Fitting Insurance Claim Loss and Claim Count Data: A Parametric View

Submission: June 29, 2018

Abstract

Actuarial loss models are of great interest for non-life insurance pricing. Though the area is studied extensively in academia, there lacks a general software implementation with a broad scope. This paper intends to summarize known results and derive analytic expressions for distributions transformed by coverage modifications. With full consideration of many frequently observed features in policy configurations, this paper can serve as a comprehensive reference for future software development. The author is currently working on an implementation in `R` based on the `actuar` and `fitdistrplus` packages. The roll-out is tentatively planned in late 2018 summer. An interactive web-based demonstration powered by `R Shiny` shall come shortly afterward.

Key Words. Coverage Modification; Transformation of Random Variables; Compound Frequency Model; Distribution Fitting for Random Losses; Parameter Estimation; Model Selection.

Contents

List of Figures

List of Tables

1 Introduction and Motivation

The actuarial profession has long been focusing on statistical analysis of insurance claim data. To date, there are many well-established methods on distribution fitting in the insurance context. (For example, [Klugman et al., 2012] is a standard reference in the industry.) However, there are very few, if any, generally applicable free softwares to fit a broad range of parametric distributions based on claim loss amount and claim count data. This project aims to consolidate relevant results, derive closed-form expressions in actuarial loss models, and guide future software development based on the framework.

The rest of the paper is organized as follows: Section 2 provides a brief overview of the mechanisms of insurance policies; Section 3 discusses theoretical results and the data fitting procedure for *severity random variables* (i.e. loss amount) under coverage modifications; Section 4 discusses theoretical results and the data fitting procedure for *frequency random variables* (i.e. loss count) under coverage modifications; Section 5 concludes the paper and provides avenues for future work. A general application in R is still under development. However, both section 3 and 4 contain numerical examples based on currently available partial implementations.

2 Definition of Coverage Modifications

2.1 Insurance Policy: From Incurred Losses to Observed Payments

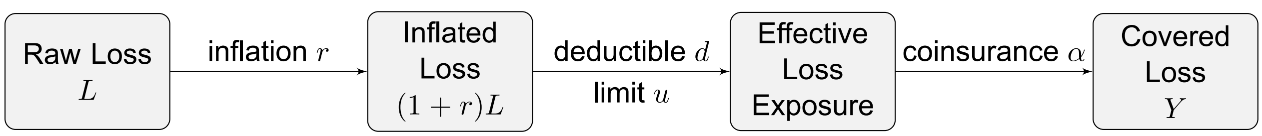
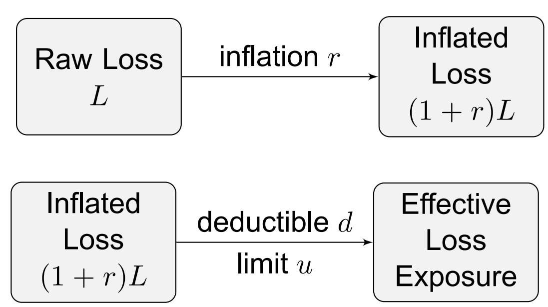


Figure 1: Mechanism of Coverage Modification

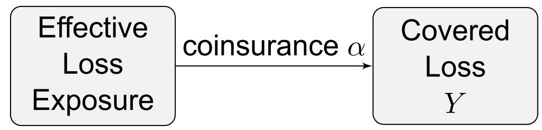
Take an auto insurance policy as an example. The policy will typically specify *deductible* amount d, *limit* amount u, and a coinsurance factor $\alpha \in (0, 1]$. Besides, if we model the loss amount now as L when the policy is written, then for the same loss, it may value more than L by the time when the loss is incurred. Therefore, actuaries also assume a constant inflation rate r between the horizon. (Note that r does not need to be positive. If $r < 0$, it simply implies a deflation rather than an inflation. However, $r > -1$ must hold because it makes no sense to have a negative loss amount.)



The *coverage modification* flow starts with a loss incurred by the policyholder. (We suppose the loss is within the range of loss types specified in the policy.) The loss $(1 + r)L$ is inflated by r with respect to our model benchmark L.

Then, loss amount $(1 + r)L$ will be considered for deductibles and limits simultaneously. The limit amount u specifies a ceiling for loss amount under consideration, i.e. any excess amount will be disregarded. There are two types of deductibles: "ordinary deductible" and "franchise deductible". For an *ordinary deductible*, the deductible amount is subtracted from the loss incurred and the insurer only pays the resulting amount if it is greater than \$ 0. On the contrary, a *franchise deductible* requires the insurer to fully reimburse the entire loss incurred when it exceeds the deductible d. In essence, given the claim is paid, a policy with franchise deductible will pay d more than one with ordinary deductible.

Finally, the loss exposure will be multiplied by a *coinsurance factor* α meaning that only a fraction of the eligible loss is covered. This feature is not common in auto policies (where $\alpha = 1$) but can appear in many other property and casualty policies.



Combining the above, the loss amounts after coverage modifications are

$$Y^L \overset{\text{def}}{=} \alpha\big[(1+r)L \wedge u - (1+r)L \wedge d\big]$$
$$\widetilde{Y}^L \overset{\text{def}}{=} \alpha\big((1+r)L \wedge u\big) \cdot \mathbb{1}\{(1+r)L > d\}$$

for a policy with ordinary deductible and franchise deductible respectively. We will see these expressions later in sections 3.7.1 and 3.7.3.

2.2 Terminology and List of Symbols

In this section, we will introduce variables and notations used throughout this paper.

Term	**Definition**
X, N, M	Some *generic random variable* to state theorems and general results.
L	*Raw size-of-loss random variable* (aka group-up loss amount) whose distribution is of interest. We assume $L \in [\,0, +\infty)$.
Y	*Modified loss random variable* from which observed data $\mathbf{Y}^n = \{y_i\}_{i=1}^n$ is generated. In most occasions, it directly relates to the claim amount paid by insurer. In the discussion, some variations of this symbol are used: a $\cdots^L$ superscript denotes *cost-per-loss* while $\cdots^P$ superscript denotes *cost-per-payment*; a $\widetilde{\ }$ sign indicates franchise deductible while without the sign indicates ordinary deductible.
N^L	*Raw loss count random variable*.
N^P	*Coverage modified loss count* (due to deductible).
d	*Policy deductible*. Can be either ordinary deductible or franchise deductible.
u	*Policy limit*.
r	Constant *inflation rate*. We assume $r > -1$.
α	*Coinsurance factor*.
Truncation	We have no information on values beyond boundaries, i.e. they cannot be detected.
Censoring	We have partial information on values beyond boundaries, i.e. they can be detected with noise, or in other words, measurements of such values are not completely known.

Table 1: List of Symbols

We will also introduce some actuarial notations. The *wedge symbol* is commonly used to denote maximum and minimum between two real numbers in actuarial science,

$$a \wedge b \overset{\text{def}}{=} \min\{a, b\} \qquad \forall a, b \in \mathbb{R} \tag{1}$$

$$a \vee b \overset{\text{def}}{=} \max\{a, b\} \qquad \forall a, b \in \mathbb{R} \tag{2}$$

In particular, for any real number $x \in \mathbb{R}$, we write

$$(x)_+ \overset{\text{def}}{=} \max\{x, 0\} \tag{3}$$

3 Coverage Modification of the Severity Random Variable

In this chapter, we start by tackling each policy feature individually and combine all of them together in section 3.7. Before diving into the analysis, we will introduce techniques for later derivation in section 3.1. Since the ideas are essentially the same, we will only state results in following sections. Note that all basic building blocks of this chapter (3.2, 3.3, 3.4, 3.5, 3.6) are covered in [Klugman et al., 2012] while partial results of section 3.7 are covered in [Dutang et al., 2008]. This chapter not only combines two references but also extend to more details and all potential policy feature combinations.

3.1 Mathematical Toolbox

3.1.1 Basic Probability Theory Arguments

Assume we have a probabilistic model for raw loss L. The key to derive the resulting model for coverage-modified loss Y is via *transformation of random variables*, i.e. we want to express

$$F_Y(y) \overset{\text{def}}{=} \mathbb{P}(Y \leq y)$$

into a form of probability involving only L and parameters. With the *cumulative distribution function* (CDF) of Y, we can derive the following associated functions on appropriate intervals,

Term	Expression
Survival Function	$S_Y(y) = 1 - F_Y(y)$
PDF: *Probability Density Function*	$f_Y(y) = \frac{d}{dy} F_Y(y)$
Hazard Function	$h_Y(y) = \frac{f_Y(y)}{S_Y(y)}$

3.1.2 Other Useful Relations

Based on the definition of the wedge symbol, we can easily obtain the following two properties,

- *Decomposition Identity*: for any random variable X and $\forall d \in \mathbb{R}$,

$$X \equiv (X \wedge d) + (X - d)_+$$

- *Linearity*: for $\forall x, y, b \in \mathbb{R}$ and $\forall a > 0$,

$$(ax+b) \wedge (ay+b) = a(x \wedge y) + b$$
$$(ax+b) \vee (ay+b) = a(x \vee y) + b$$

Furthermore, they are two more handy results to transform the expression of expectations. Let X denote a random variable and consider $\forall a > 0$ and $\forall b \in \mathbb{R}$,

$$\mathbb{E}[aX \wedge b] = a\mathbb{E}\big[X \wedge \frac{b}{a}\big]$$
$$\mathbb{E}[(aX - b)_+] = a\big(\mathbb{E}[X] - \mathbb{E}\big[X \wedge \frac{b}{a}\big]\big)$$

3.2 Ordinary Policy Deductible

3.2.1 Cost-Per-Loss: Left-Censored-and-Shifted Random Variable

$$Y^L \overset{\text{def}}{=} (L-d)_+ = \begin{cases} 0 & \text{, if } L \leq d \\ L-d & \text{, if } L > d \end{cases} \tag{4}$$

For abbreviation, let Y denote Y^L in stating the following results.

- Y has support $[0, +\infty)$
- For any $y \in [0, +\infty)$, the CDF and survival functions are given by

$$F_Y(y) = F_L(y+d) \qquad S_Y(y) = S_L(y+d)$$

- Y has a mixed distribution even if L is continuous due to mass at zero
  - Probability mass at $y = 0$ is given by

$$p_Y(0) \overset{\text{def}}{=} \mathbb{P}(Y=0) = \mathbb{P}(L \leq d) = F_L(d)$$

  - For any $y \in (0, +\infty)$, the PDF is given by

$$f_Y(y) = f_L(y+d)$$

    and the hazard function can be derived on the same support

$$h_Y(y) = h_L(y+d)$$

- The expectation is linked to L and d via

$$\mathbb{E}[Y] = \mathbb{E}[L] - \mathbb{E}[L \wedge d]$$

Y^L is also known as *cost-per-loss*, i.e. claim amount paid per loss event.

3.2.2 Cost-Per-Payment: Left-Truncated-and-Shifted Random Variable

$$Y^P \overset{\text{def}}{=} (L-d) \mid L > d = \begin{cases} \text{undefined} & \text{, if } L \le d \\ L-d & \text{, if } L > d \end{cases} \tag{5}$$

For abbreviation, let Y denote Y^P in stating the following results.

- Y has support $(0, +\infty)$
- For any $y \in (0, +\infty)$, the CDF and survival functions are given by

$$F_Y(y) = \frac{F_L(y+d) - F_L(d)}{1 - F_L(d)} \qquad S_Y(y) = \frac{S_L(y+d)}{S_L(y)}$$

- Y will be continuous if L is continuous, and its resulting PDF on $y \in (0, +\infty)$ is given by

$$f_Y(y) = \frac{f_L(y+d)}{1 - F_L(d)}$$

  and the hazard function can be derived on the same support

$$h_Y(y) = h_L(y+d)$$

- The expectation is linked to L and d via

$$\mathbb{E}[Y] = \frac{\mathbb{E}[L] - \mathbb{E}[L \wedge d]}{1 - F_L(d)}$$

Y^P is also known as *cost-per-payment*, i.e. claim amount paid per payment event.

3.3 Franchise Policy Deductibles

3.3.1 Cost-Per-Loss: Left-Censored Random Variable

$$\widetilde{Y}^L \overset{\text{def}}{=} L \cdot \mathbb{1}\{L > d\} = \begin{cases} 0 & \text{, if } L \le d \\ L & \text{, if } L > d \end{cases} \tag{6}$$

For abbreviation, let Y denote $\widetilde{Y}^L$ in stating the following results.

- Y has support $\{0\} \cup (d, +\infty)$
- For any $y \in [0, +\infty)$, the CDF and survival functions are given by

$$F_Y(y) = \begin{cases} F_L(d) & \text{, if } 0 \le y \le d \\ F_L(y) & \text{, if } d < y \end{cases} \qquad S_Y(y) = \begin{cases} S_L(d) & \text{, if } 0 \le y \le d \\ S_L(y) & \text{, if } d < y \end{cases}$$

- Y has a mixed distribution even if L is continuous due to discontinuity in the support
  - Probability mass at $y = 0$ is given by

$$p_Y(0) \overset{\text{def}}{=} \mathbb{P}(Y = 0) = \mathbb{P}(L \le d) = F_L(d)$$

– For any $y \in (d, +\infty)$, the PDF is given by

$$f_Y(y) = f_L(y)$$

and the hazard function can be derived on the same support

$$h_Y(y) = h_L(y)$$

- The expectation is linked to L and d via

$$\mathbb{E}[Y] = \mathbb{E}[L] - \mathbb{E}[L \wedge d] + d\big(1 - F_L(d)\big)$$

3.3.2 Cost-Per-Payment: Left-Truncated Random Variable

$$\widetilde{Y}^P \overset{\text{def}}{=} L \mid L > d = \begin{cases} \text{undefined} & \text{, if } L \leq d \\ L & \text{, if } L > d \end{cases} \tag{7}$$

For abbreviation, let Y denote $\widetilde{Y}^P$ in stating the following results.

- Y has support $[d, +\infty)$
- For any $y \in [0, +\infty)$, the CDF and survival functions are given by

$$F_Y(y) = \begin{cases} 0 & \text{, if } 0 \leq y < d \\ \frac{F_L(y) - F_L(d)}{1 - F_L(d)} & \text{, if } d < y \end{cases} \qquad S_Y(y) = \begin{cases} 1 & \text{, if } 0 \leq y < d \\ \frac{S_L(y)}{S_L(d)} & \text{, if } d < y \end{cases}$$

- Y will be continuous if L is continuous, and its resulting PDF on $y \in (d, +\infty)$ is given by

$$f_Y(y) = \frac{f_L(y)}{1 - F_L(d)}$$

  and the hazard function can be derived on the same support

$$h_Y(y) = h_L(y)$$

- The expectation is linked to L and d via

$$\mathbb{E}[Y] = \frac{\mathbb{E}[L] - \mathbb{E}[L \wedge d]}{1 - F_L(d)} + d$$

3.4 Policy Limits: Right Censored Random Variable

$$Y \overset{\text{def}}{=} L \wedge u = \begin{cases} L & \text{, if } L < u \\ u & \text{, if } L \geq u \end{cases} \tag{8}$$

Similarly, we present the following properties for Y.

- Y has support $[0, u]$
- For any $y \in [0, +\infty)$, the CDF and survival functions are given by

$$F_Y(y) = \begin{cases} F_L(y) & \text{, if } 0 \leq y < u \\ 1 & \text{, if } u \leq y \end{cases} \qquad S_Y(y) = \begin{cases} S_L(y) & \text{, if } 0 \leq y < u \\ 0 & \text{, if } u \leq y \end{cases}$$

- Y has a mixed distribution even if L is continuous due to mass at u
  - Probability mass at $y = u$ is given by

$$p_Y(u) = \mathbb{P}(Y = u) = 1 - F_L(u)$$

  - For any $y \in [0, u)$, its PDF is given by

$$f_Y(y) = f_L(y)$$

    and the hazard function can be derived on the same support

$$h_Y(y) = h_L(y)$$

- The expectation is linked to L and d via

$$\mathbb{E}[Y] = \mathbb{E}[L \wedge u]$$

3.5 Constant Inflation

Let r denote the constant *inflation rate*, then we can define

$$Y \overset{\text{def}}{=} (1 + r)L \tag{9}$$

Inflated loss Y has the following properties.

- Y has support $[0, +\infty)$
- For any $y \in [0, +\infty)$, the CDF and survival functions are given by

$$F_Y(y) = F_L(\frac{y}{1+r}) \qquad S_Y(y) = S_L(\frac{y}{1+r})$$

- Y will be continuous if L is continuous, and its resulting PDF on $y \in (0, +\infty)$ is given by

$$f_Y(y) = \frac{1}{1+r} f_L(\frac{y}{1+r})$$

  and the hazard function can be derived on the same support

$$h_Y(y) = \frac{1}{1+r} h_L(\frac{y}{1+r})$$

- The expectation is linked to L and d via

$$\mathbb{E}[Y] = (1 + r)\mathbb{E}[L]$$

3.6 Coinsurance

In a policy with a *coinsurance factor* $\alpha \in (0, 1]$, the insurer's portion of the loss is

$$Z \overset{\text{def}}{=} \alpha Y \tag{10}$$

By convention, the coinsurance factor is applied to modified loss random variable Y which has already considered deductibles, limits, and inflation with respect to the underlying raw random variable L. Hence, we use another letter Z to distinguish from Y. Modified loss random variable after coinsurance Z has the following properties.

- Z's support is α times Y's support
- For any $z \in [0, +\infty)$, the CDF and survival functions are given by

$$F_Z(z) = F_Y(\frac{z}{\alpha}) \qquad S_Z(z) = S_Y(\frac{z}{\alpha})$$

- Z will be continuous if Y is continuous, and its resulting PDF on an open interval in its support is given by

$$f_Z(z) = \frac{1}{\alpha} f_Y(\frac{z}{\alpha})$$

  and the hazard function can be derived on the same support

$$h_Z(z) = \frac{1}{\alpha} h_Y(\frac{z}{\alpha})$$

3.7 Combinations of Deductible, Limit, Inflation, and Coinsurance

In this section, we combine results from previous sections to obtain the probability model of modified loss random variables. Below is a brief overview of all possible scenarios.

Subsection Title	Setup	
	Deductible	Modified Loss
Shifted Left Censoring and Right Censoring	Ordinary	Cost-per-loss
Shifted Left Truncation and Right Censoring	Ordinary	Cost-per-payment
Left Censoring and Right Censoring	Franchise	Cost-per-loss
Left Truncation and Right Censoring	Franchise	Cost-per-payment

3.7.1 Shifted Left Censoring and Right Censoring

Policy Configuration	Notation
Ordinary Deductible	d
Policy Limit	$u(> d)$
Constant Inflation Rate	r
Coinsurance Factor	α

We consider the cost-per-loss random variable Y^L, which is defined as follows,

$$Y^L \overset{\text{def}}{=} \alpha\big[(1+r)L \wedge u - (1+r)L \wedge d\big] \tag{11}$$

$$= \begin{cases} 0 & \text{, if} \quad 0 \le L \le \dfrac{d}{1+r} \\ \alpha((1+r)L - d) & \text{, if} \quad \dfrac{d}{1+r} < L \le \dfrac{u}{1+r} \\ \alpha(u-d) & \text{, if} \quad \dfrac{u}{1+r} < L \end{cases} \tag{12}$$

For abbreviation, let Y denote Y^L in stating the following results.

- Y has support $[0, \alpha(u-d)]$
- For any $y \in [0, +\infty)$, the CDF and survival functions are given by

$$F_Y(y) = \begin{cases} F_L(\dfrac{d}{1+r}) & \text{, if} \quad 0 = y \\ F_L(\dfrac{y}{\alpha(1+r)} + \dfrac{d}{1+r}) & \text{, if} \quad 0 < y < \alpha(u-d) \\ 1 & \text{, if} \quad \alpha(u-d) \le y \end{cases}$$

$$S_Y(y) = \begin{cases} S_L(\dfrac{d}{1+r}) & \text{, if} \quad 0 = y \\ S_L(\dfrac{y}{\alpha(1+r)} + \dfrac{d}{1+r}) & \text{, if} \quad 0 < y < \alpha(u-d) \\ 0 & \text{, if} \quad \alpha(u-d) \le y \end{cases}$$

- For any $p \in [0, 1]$, the quantile function is given by

$$Q_Y(p) = \begin{cases} 0 & \text{, if} \quad 0 \le p \le F_L(\dfrac{d}{1+r}) \\ \alpha(1+r)Q_L(p) - \alpha d & \text{, if} \quad F_L(\dfrac{d}{1+r}) < p \le F_L(\dfrac{u}{1+r}) \\ \alpha(u-d) & \text{, if} \quad F_L(\dfrac{u}{1+r}) < p \le 1 \end{cases}$$

- Y has a mixed distribution even if L is continuous due to mass at end points $y = 0$ and $y = \alpha(u-d)$. However, Y will be continuous in the open interval $\big(0, \alpha(u-d)\big)$.
  - Probability mass at $y = 0$ is given by

$$p_Y(0) = \mathbb{P}(L \le \frac{d}{1+r}) = F_L(\frac{d}{1+r})$$

  - For any $y \in \big(0, \alpha(u-d)\big)$, the PDF is given by

$$f_Y(y) = \frac{1}{\alpha(1+r)} f_L(\frac{y}{\alpha(1+r)} + \frac{d}{1+r})$$

    and the hazard function can be derived on the same support

$$h_Y(y) = \frac{1}{\alpha(1+r)} h_L(\frac{y}{\alpha(1+r)} + \frac{d}{1+r})$$

– Probability mass at $y = \alpha(u-d)$ is given by

$$p_Y(\alpha(u-d)) = \mathbb{P}(L > \frac{u}{1+r}) = 1 - F_L(\frac{u}{1+r})$$

- The first two moments are linked to L, u, and d via

$$\mathbb{E}[Y] = \alpha(1+r)\big(\mathbb{E}[L \wedge u^*] - \mathbb{E}[L \wedge d^*]\big)$$
$$\mathbb{E}[Y^2] = \alpha^2(1+r)^2\Big(\mathbb{E}[(L \wedge u^*)^2] - \mathbb{E}[(L \wedge d^*)^2] - 2d^*\big(\mathbb{E}[(L \wedge u^*)] - \mathbb{E}[L \wedge d^*]\big)\Big)$$

where $u^* = \frac{u}{1+r}$ and $d^* = \frac{d}{1+r}$.

3.7.2 Shifted Left Truncation and Right Censoring

Policy Configuration	Notation
Ordinary Deductible	d
Policy Limit	$u(> d)$
Constant Inflation Rate	r
Coinsurance Factor	α

We consider the cost-per-payment random variable Y^P, which is defined as follows,

$$Y^P \stackrel{\text{def}}{=} \alpha\big[(1+r)L \wedge u - d\big] \;\big|\; (1+r)L > d \tag{13}$$
$$= \begin{cases} \text{undefined} & \text{, if} \quad 0 \le L \le \dfrac{d}{1+r} \\ \alpha((1+r)L - d) & \text{, if} \quad \dfrac{d}{1+r} < L \le \dfrac{u}{1+r} \\ \alpha(u-d) & \text{, if} \quad \dfrac{u}{1+r} < L \end{cases} \tag{14}$$

For abbreviation, let Y denote Y^P in stating the following results.

- Y has support $(0, \alpha(u-d)]$
- For any $y \in [0, +\infty)$, the CDF and survival functions are given by

$$F_Y(y) = \begin{cases} 0 & \text{, if} \quad 0 = y \\ \dfrac{F_L(\frac{y}{\alpha(1+r)} + \frac{d}{1+r}) - F_L(\frac{d}{1+r})}{1 - F_L(\frac{d}{1+r})} & \text{, if} \quad 0 < y < \alpha(u-d) \\ 1 & \text{, if} \quad \alpha(u-d) \le y \end{cases}$$
$$S_Y(y) = \begin{cases} 1 & \text{, if} \quad 0 = y \\ \dfrac{S_L(\frac{y}{\alpha(1+r)} + \frac{d}{1+r})}{S_L(\frac{d}{1+r})} & \text{, if} \quad 0 < y < \alpha(u-d) \\ 0 & \text{, if} \quad \alpha(u-d) \le y \end{cases}$$

- For any $p \in [0,1]$, the quantile function is given by

$$Q_Y(p) = \begin{cases} \alpha(1+r)Q_L\Big(\big(1-F_L(\frac{d}{1+r})\big)p + F_L(\frac{d}{1+r})\Big) - \alpha d & \text{, if } p \in \left[0, \frac{F_L(\frac{u}{1+r}) - F_L(\frac{d}{1+r})}{1-F_L(\frac{d}{1+r})}\right] \\ \alpha(u-d) & \text{, if } p \in \left[\frac{F_L(\frac{u}{1+r}) - F_L(\frac{d}{1+r})}{1-F_L(\frac{d}{1+r})}, 1\right] \end{cases}$$

- Y has a mixed distribution even if L is continuous due to mass at end point $y = \alpha(u-d)$. However, Y will be continuous in the open interval $\big(0, \alpha(u-d)\big)$.
  - For any $y \in \big(0, \alpha(u-d)\big)$, the PDF is given by

$$f_Y(y) = \frac{1}{\alpha(1+r)} \frac{f_L(\frac{y}{\alpha(1+r)} + \frac{d}{1+r})}{1-F_L(\frac{d}{1+r})}$$

    and the hazard function can be derived on the same support

$$h_Y(y) = \frac{1}{\alpha(1+r)} h_L(\frac{y}{\alpha(1+r)} + \frac{d}{1+r})$$

  - Probability mass at $y = \alpha(u-d)$ is given by

$$p_Y(\alpha(u-d)) = \mathbb{P}\left(L > \frac{u}{1+r} \middle| L > \frac{d}{1+r}\right) = \frac{1-F_L(\frac{u}{1+r})}{1-F_L(\frac{d}{1+r})}$$

- The first two moments are linked to L, u, and d via

$$\mathbb{E}[Y] = \frac{\alpha(1+r)}{1-F_L(d^*)}\big(\mathbb{E}[L \wedge u^*] - \mathbb{E}[L \wedge d^*]\big)$$
$$\mathbb{E}[Y^2] = \frac{\alpha^2(1+r)^2}{1-F_L(d^*)}\Big(\mathbb{E}[(L \wedge u^*)^2] - \mathbb{E}[(L \wedge d^*)^2] - 2d^*\big(\mathbb{E}[(L \wedge u^*)] - \mathbb{E}[L \wedge d^*]\big)\Big)$$

  where $u^* = \frac{u}{1+r}$ and $d^* = \frac{d}{1+r}$.

3.7.3 Left Censoring and Right Censoring

Policy Configuration	Notation
Franchise Deductible	d
Policy Limit	$u(> d)$
Constant Inflation Rate	r
Coinsurance Factor	α

We consider the cost-per-loss random variable $\widetilde{Y}^L$, which is defined as follows,

$$\widetilde{Y}^L \overset{\text{def}}{=} \alpha\big((1+r)L \wedge u\big) \cdot \mathbb{1}\{(1+r)L > d\} \tag{15}$$

$$= \begin{cases} 0 & \text{, if} \quad 0 \le L \le \frac{d}{1+r} \\ \alpha(1+r)L & \text{, if} \quad \frac{d}{1+r} < L \le \frac{u}{1+r} \\ \alpha u & \text{, if} \quad \frac{u}{1+r} < L \end{cases} \tag{16}$$

For abbreviation, let Y denote $\widetilde{Y}^L$ in stating the following results.

- Y has support $\{0\} \cup (\alpha d, \alpha u]$
- For any $y \in [0, +\infty)$, the CDF function is given by

$$F_Y(y) = \begin{cases} F_L(\dfrac{d}{1+r}) & \text{, if} \quad 0 \le y \le \alpha d \\ F_L(\dfrac{y}{\alpha(1+r)}) & \text{, if} \quad \alpha d < y < \alpha u \\ 1 & \text{, if} \quad \alpha u \le y \end{cases} \qquad S_Y(y) = \begin{cases} S_L(\dfrac{d}{1+r}) & \text{, if} \quad 0 \le y \le \alpha d \\ S_L(\dfrac{y}{\alpha(1+r)}) & \text{, if} \quad \alpha d < y < \alpha u \\ 0 & \text{, if} \quad \alpha u \le y \end{cases}$$

- For any $p \in [0, 1]$, the quantile function is given by

$$Q_Y(p) = \begin{cases} 0 & \text{, if} \quad 0 \le p \le F_L(\dfrac{d}{1+r}) \\ \alpha(1+r)Q_L(p) & \text{, if} \quad F_L(\dfrac{d}{1+r}) < p \le F_L(\dfrac{u}{1+r}) \\ \alpha u & \text{, if} \quad F_L(\dfrac{u}{1+r}) < p \le 1 \end{cases}$$

- Y has a mixed distribution even if L is continuous due to mass at end points $y = 0$ and $y = \alpha u$. However, Y will be continuous in the open interval $(\alpha d, \alpha u)$.
  - Probability mass at $y = 0$ is given by

$$p_Y(0) = \mathbb{P}(L \le \frac{d}{1+r}) = F_L(\frac{d}{1+r})$$

  - For any $y \in (\alpha d, \alpha u)$, the PDF is given by

$$f_Y(y) = \frac{1}{\alpha(1+r)} f_L(\frac{y}{\alpha(1+r)})$$

    and the hazard function can be derived on the same support

$$h_Y(y) = \frac{1}{\alpha(1+r)} h_L(\frac{y}{\alpha(1+r)})$$

  - Probability mass at $y = \alpha(u - d)$ is given by

$$p_Y(\alpha u) = \mathbb{P}(L > \frac{u}{1+r}) = 1 - F_L(\frac{u}{1+r})$$

- The k-th moment is linked to L, u, and d via

$$\mathbb{E}[Y^k] = \alpha^k(1+r)^k\big(\mathbb{E}[(L \wedge u^*)^k] - \mathbb{E}[(L \wedge d^*)^k]\big) + \alpha^k d^k\big(1 - F_L(d^*)\big)$$

  where $u^* = \frac{u}{1+r}$ and $d^* = \frac{d}{1+r}$.

3.7.4 Left Truncation and Right Censoring

Policy Configuration	Notation
Franchise Deductible	d
Policy Limit	$u(> d)$
Constant Inflation Rate	r
Coinsurance Factor	α

We consider the cost-per-payment random variable $\widetilde{Y}^P$, which is defined as follows,

$$\widetilde{Y}^P \overset{\text{def}}{=} \alpha\big((1+r)L \wedge u\big)\big|(1+r)L > d \tag{17}$$

$$= \begin{cases} \text{undefined} & \text{, if} \quad 0 \le L \le \dfrac{d}{1+r} \\ \alpha(1+r)L & \text{, if} \quad \dfrac{d}{1+r} < L \le \dfrac{u}{1+r} \\ \alpha u & \text{, if} \quad \dfrac{u}{1+r} < L \end{cases} \tag{18}$$

For abbreviation, let Y denote $\widetilde{Y}^P$ in stating the following results.

- Y has support $(\alpha d, \alpha u]$
- For any $y \in [0, +\infty)$, the CDF and survival functions are given by

$$F_Y(y) = \begin{cases} 0 & \text{, if} \quad 0 \le y \le \alpha d \\ \dfrac{F_L(\frac{y}{\alpha(1+r)}) - F_L(\frac{d}{1+r})}{1 - F_L(\frac{d}{1+r})} & \text{, if} \quad \alpha d < y < \alpha u \\ 1 & \text{, if} \quad \alpha u \le y \end{cases} \qquad S_Y(y) = \begin{cases} 1 & \text{, if} \quad 0 \le y \le \alpha d \\ \dfrac{S_L(\frac{y}{\alpha(1+r)})}{S_L(\frac{d}{1+r})} & \text{, if} \quad \alpha d < y < \alpha u \\ 0 & \text{, if} \quad \alpha u \le y \end{cases}$$

- For any $p \in [0, 1]$, the quantile function is given by

$$Q_Y(p) = \begin{cases} \alpha(1+r)Q_L\Big(\big(1 - F_L(\dfrac{d}{1+r})\big)p + F_L(\dfrac{d}{1+r})\Big) & \text{, if } p \in \left[0, \dfrac{F_L(\frac{u}{1+r}) - F_L(\frac{d}{1+r})}{1 - F_L(\frac{d}{1+r})}\right] \\ \alpha u & \text{, if } p \in \left[\dfrac{F_L(\frac{u}{1+r}) - F_L(\frac{d}{1+r})}{1 - F_L(\frac{d}{1+r})}, 1\right] \end{cases}$$

- Y has a mixed distribution even if L is continuous due to mass at end point $y = \alpha u$. However, Y will be continuous in the open interval $(\alpha d, \alpha u)$.
  - For any $y \in (\alpha d, \alpha u)$, the PDF is given by

  $$f_Y(y) = \frac{1}{\alpha(1+r)} \frac{f_L(\frac{y}{\alpha(1+r)})}{1 - F_L(\frac{d}{1+r})}$$

  and the hazard function can be derived on the same support

  $$h_Y(y) = \frac{1}{\alpha(1+r)} h_L(\frac{y}{\alpha(1+r)})$$

– Probability mass at $y = \alpha(u - d)$ is given by

$$p_Y(\alpha u) = \mathbb{P}\left(L > \frac{u}{1+r} \middle| L > \frac{d}{1+r}\right) = \frac{1 - F_L(\frac{u}{1+r})}{1 - F_L(\frac{d}{1+r})}$$

- The k-th moment is linked to L, u, and d via

$$\mathbb{E}[Y^k] = \frac{\alpha^k(1+r)^k}{1 - F_L(d^*)}\big(\mathbb{E}[(L \wedge u^*)^k] - \mathbb{E}[(L \wedge d^*)^k]\big) + \alpha^k d^k$$

  where $u^* = \frac{u}{1+r}$ and $d^* = \frac{d}{1+r}$.

3.8 General Patterns

Many results derived in this chapter seem rather complicated. The purpose of this section is to shed some light upon intuitions and relations between results.

3.8.1 Limited Expected Value Function

Notice that we simplified moments to forms involving

$$\mathbb{E}\big[(L \wedge l)^k\big]$$

which is clearly a function of order k and limit value l once the distribution of L is specified. This expression is known as the *limited expected value function* of random variable L. It has many desirable properties and has wide applications in actuarial science. Closed-form expressions are available for almost all commonly used parametric families in the Appendix of [Klugman et al., 2012].

3.8.2 Cost-Per-Loss and Cost-Per-Payment

By definition, the insurer will pay if the covered loss is positive, i.e.

$$Y^P \overset{\text{def}}{=} Y^L \big| Y^L > 0 \tag{19}$$

Taking expectation on both sides, we get

$$\mathbb{E}[Y^P] = \frac{\mathbb{E}[Y^L]}{1 - F_L(d^*)}$$

where $d^* = \frac{d}{1+r}$. Note that the above relations hold for both ordinary and franchise deductibles as long as two sides match.

3.8.3 Ordinary Deductible and Franchise Deductible

In 2.1, we concluded that “given the claim is paid, a policy with franchise deductible will pay d more than one with ordinary deductible. ” Translating to an equation, we have

$$\widetilde{Y} = Y + \alpha d \cdot \mathbb{1}\{Y > 0\}$$

Taking expectation on both sides, we get

$$\mathbb{E}[\widetilde{Y}] = \mathbb{E}[Y] + \alpha d\big(1 - F_L(d^*)\big)$$

where $d^* = \frac{d}{1+r}$. Note that both $\cdots^L$ and $\cdots^P$ superscripts satisfy this relation as long as two sides match.

3.8.4 “Effective” Parameters

Under the full (d, u, r, α) specification, we hope to re-parametrize them to obtain some insights. Based on the two above results in this section, it suffices to study the case of Y^L, i.e. the cost-per-loss random variable with an ordinary deductible. From 3.7.1, we know that it has definition

$$Y^L \overset{\text{def}}{=} \alpha\big[(1+r)L \wedge u - (1+r)L \wedge d\big] = \alpha(1+r)\left[L \wedge \frac{u}{1+r} - L \wedge \frac{d}{1+r}\right]$$

Therefore, we can define the following parametrization

$$\alpha^* \overset{\text{def}}{=} \alpha(1+r) \qquad u^* \overset{\text{def}}{=} \frac{u}{1+r} \qquad d^* \overset{\text{def}}{=} \frac{d}{1+r} \qquad \widetilde{d} \overset{\text{def}}{=} \alpha d$$

Then, we can re-write the definitions in sections 3.7.1 and 3.7.2 into

$$Y^L = \alpha^*\big[L \wedge u^* - L \wedge d^*\big] \qquad \widetilde{Y}^L = \alpha^*\big(L \wedge u^*\big) \cdot \mathbb{1}\{L > d^*\}$$

and the resulting expressions for sections 3.7.3 and 3.7.4 can be expressed in terms of conditions.

3.9 Data Fitting Procedure

Suppose we have a parametric probabilistic model for raw loss size L. Given policy specification

$$(u, d, r, \alpha, \text{deductible type: ordinary/franchise, cost-per-loss/payment})$$

we can work out the distribution characteristics of Y, namely F_Y, Q_y, and f_Y. If we observe data for coverage modified losses Y of sample size n, denoted by $\mathbf{Y}^n = \{y_i\}_{i=1}^n$, then we can fit such data on our parametric model for Y using any acceptable method such as *Maximum Likelihood Estimation* (MLE), *Moments Matching Estimation* (MME), and *Percentile Matching Estimation* (PME). However, notice that our closed-form expressions are only available up to the second-order moment for policies with ordinary deductibles. In such case, if one insists on using MME beyond the second-order term, one may need to consider numerical quadrature routines.

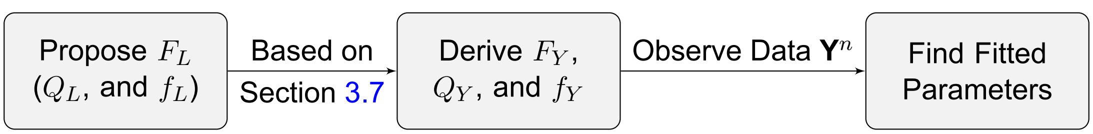


Figure 2: Data Fitting Procedure for Coverage Modified Severity Random Variable

3.10 Numerical Example

In this section, we simulate data of size $n = 100$ given the following policy configurations: ordinary deductible $d = \$5,000$, policy limit $u = \$20,000$, coinsurance factor $\alpha = 0.9$, under a constant inflation rate $r = 5\%$, and with a log-normal raw loss size $\ln L \sim \mathcal{N}(\mu = 9, \sigma^2 = 1)$. Consider the cost-per-loss random variable. We plot simulated data below for both raw loss and modified loss with their respective theoretical PDFs.

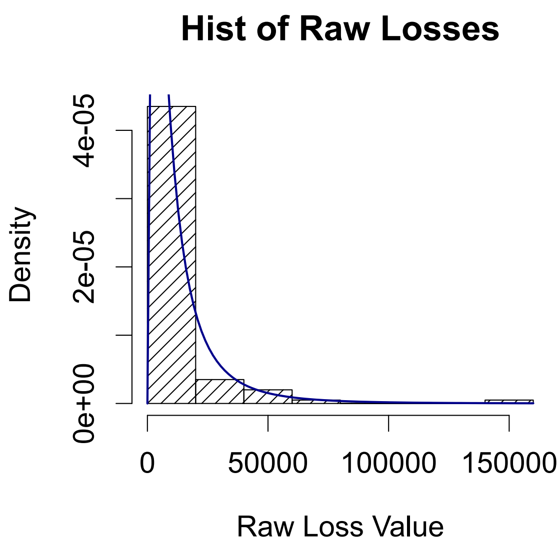


Figure 3: Example 3.10 Raw Loss Size

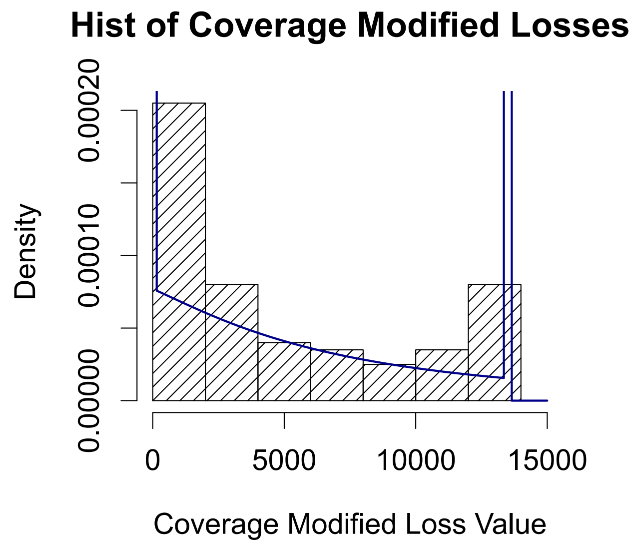


Figure 4: Example 3.10 Modified Loss Size

Given the above specification and simulated data, we can estimate parameters (μ, σ^2) using MLE, MME with first and second moments, and PME with 33% and 66% percentiles. In addition, we perform bootstrap for $B = 1000$ times for each method to estimate the standard errors of the estimations. A summary is listed below.

	Maximum Likelihood			Moments Matching			Quantile Matching		
	Median	2.5%	97.5%	Median	2.5%	97.5%	Median	2.5%	97.5%
μ	8.9807	8.8002	9.1526	7.7807	6.7830	8.3564	9.0208	8.5113	9.1897
σ	0.8271	0.6811	1.0356	2.9098	2.2703	3.6912	0.8679	0.5682	1.7104

Table 2: Example 3.10 Estimated Parameters and Non-Parametric Bootstrap Confidence Intervals

Recall that the true parameter values are $(\mu = 9, \sigma = 1)$. We can observe from Table 2 that MLE is the most accurate estimate while MME is the least accurate.

4 Coverage Modification of the Frequency Random Variable

This chapter has a similar plan as the previous chapter on severity random variables. Some statistical terminologies may not be as familiar to audience outside from actuaries, so we will provide more details in section 4.1. Besides, the probabilistic model for coverage modified loss count random variable is rather trickier to derive, but fortunately there is only one scenario to consider — deductibles, because loss count only depends on whether an incurred loss is covered/paid or not and hence does not depend on the magnitude.

4.1 Mathematical Toolbox

4.1.1 The $(a, b, 0)$ and $(a, b, 1)$ Classes of Discrete Distributions

The $(a, b, 0)$ class of (non-negative) discrete distributions satisfy the following recursive relation

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k} \qquad \forall a, b \in \mathbb{R}, k \in \mathbb{N} \tag{20}$$

where p_k denotes probability mass evaluated at k, i.e. $\mathbb{P}(N = k)$. Based on this definition, more general distributions can be defined by modifying probability mass to p_0^M at $k = 0$ but still maintain the recursive relation for $k \in \mathbb{N}$ and $k \geq 1$, i.e.

$$p_0 \to p_0^M \qquad \text{and} \qquad p_k^M = \frac{1 - p_0^M}{1 - p_0} p_k \quad k = 1, 2, 3, \cdots$$

In particular, when $p_0^M = 0$, we call it a *zero-truncated distribution*. This broader definition give rise to the $(a, b, 1)$ class. To be precise, the $(a, b, 1)$ class of (non-negative) discrete distributions satisfy the following recursive relation

$$\frac{p_k^M}{p_{k-1}^M} = a + \frac{b}{k} \qquad \forall a, b \in \mathbb{R}, k \in \mathbb{N}_+ \qquad \text{and} \qquad \forall p_0^M \in [0, 1] \tag{21}$$

It is easy to see that the $(a, b, 0)$ class is a subset of the $(a, b, 1)$ class. *Table 6.4 Members of the* $(a, b, 1)$ *class* in [Klugman et al., 2012] summarizes some commonly used distributions in the class and their respective p_0, a, b values.

4.1.2 Probability Generating Function

The *probability generating function* (*PGF*) of a (non-negative integer-valued) discrete random variable N is defined as

$$G_N(z) \overset{\text{def}}{=} \mathbb{E}[Z^N] = \sum_{n=0}^{\infty} \mathbb{P}(N = n) \cdot z^n \tag{22}$$

Similar to a *moment generating function*, this expression $G_N(z)$ can "generate" probability masses by taking derivatives,

$$\mathbb{P}(N = k) = \frac{{G_N}^{(k)}(0)}{k!}$$

Furthermore, the expectation of N is given by

$$\mathbb{E}[N] = \frac{d}{dz} G_N(z) \bigg|_{z \to 1^-} \overset{\text{denote}}{=} {G_N}'(1^-)$$

and the k-th factorial moment of N is given by

$$\mathbb{E}[N(N-1)\cdots(N-k+1)] = \mathbb{E}\left[\frac{N!}{(N-k)!}\right] = \frac{d^k}{dz^k} G_N(z) \bigg|_{z \to 1^-} \overset{\text{denote}}{=} {G_N}^{(k)}(1^-)$$

which yields the following expression for variance of N,

$$\text{Var}[N] = {G_N}''(1^-) + {G_N}'(1^-) - \left({G_N}'(1^-)\right)^2$$

4.1.3 Compound Frequency Model

Let X follows a *compound frequency model* with discrete *primary distribution* N and discrete *secondary distribution* M, i.e.

$$X \overset{\text{def}}{=} \sum_{i=1}^{N} M_i \tag{23}$$

Then, the PGF of X is given by

$$G_X(z) = G_N\big(G_M(z)\big)$$

4.2 Coverage Modified Loss Count

First, we identify the relationship between observed loss count N^P and raw loss count N^P,

$$N^P \overset{\text{def}}{=} \sum_{i=1}^{N^L} \mathbb{1}\big\{(1+r)L > d\big\} \tag{24}$$

Observe that only effective deductibles ($d^* = \frac{d}{1+r}$) will affect the observed loss count (i.e. N^P that incurs payments). In addition, both franchise deductible and ordinary deductible have the same effects. We can also notice that N^P follows a compound frequency model with primary distribution N^L and secondary distribution $\mathbb{1}\big\{(1+r)L > d\big\}$, which is equivalent to a Bernoulli distribution with success probability

$$\theta = \mathbb{P}\big(L > \frac{d}{1+r}\big)$$

Therefore, the PGF of N^P is given by

$$G_{N^P}(z) = G_{N^L}\big(G_{\mathbb{1}\{(1+r)L>d\}}(z)\big) = G_{N^L}(1-\theta+\theta z)$$

or equivalently,

$$G_{N^L}(z) = G_{N^P}\big(1 - \frac{1}{\theta} + \frac{1}{\theta} z\big)$$

We aim to simplify the data fitting procedure by finding a condition when the resulting N^P is from the same parametric family as its underlying N^L. The following result in [Klugman et al., 2012] provides a sufficient condition,

Theorem 4.1. *Suppose N^L depends on parameters (α, β) such that*

$$G_{N^L}(z;\alpha,\beta) = \alpha + (1-\alpha)\frac{B[\beta(z-1)] - B(-\beta)}{1 - B(-\beta)}$$

where $B(\cdot)$ is functionally independent of β. Then $G_{N^P}(z) = G_{N^L}(z;\alpha^*,\beta^*)$ where "modified" parameters α^*, β^* are given by

$$\alpha^* = \alpha + (1-\alpha)\frac{B(-\theta\beta) - B(-\beta)}{1 - B(-\beta)} \qquad \beta^* = \theta\beta$$

Notice that $\alpha = \mathbb{P}(N^L = 0)$ and $\alpha^* = \mathbb{P}(N^P = 0) = G_{N^L}(1-\theta;\alpha,\beta)$.

Fortunately, all members of $(a, b, 0)$ and $(a, b, 1)$ classes satisfy the above condition. Hence, we can use the same family to fit observed data of N^P and recover the parameters for N^L. A summary of parameter transformation is given below. The table is adapted from *Table 8.3 Frequency adjustments* of [Klugman et al., 2012].

Family	N^L **Parameters**	N^P **Parameters**
Poisson	λ	$\lambda^* = \theta\lambda$
Zero-Modified Poisson	λ, p_0^M	$\lambda^* = \theta\lambda$, $p_0^{M*} = \frac{1 - e^{-\lambda^*}}{1 - e^{-\lambda}} p_0^M + \frac{e^{-\lambda^*} - e^{-\lambda}}{1 - e^{-\lambda}}$
Binomial	p, with n known	$p^* = \theta p$
Zero-Modified Binomial	p, p_0^M, with n known	$p^* = \theta p$, $p_0^{M*} = \frac{1 - (1 - p^*)^n}{1 - (1 - p)^n} p_0^M + \frac{(1 - p^*)^n - (1 - p)^n}{1 - (1 - p)^n}$
Negative Binomial	r, p	$r^* = r$, $p^* = p$
Zero-Modified Negative Binomial	r, p,p_0^M	$r^* = r$, $p^* = p$, $p_0^{M*} = \frac{1 - (1 + \beta^*)^{-r}}{1 - (1 + \beta)^{-r}} p_0^M + \frac{(1 + \beta^*)^{-r} - (1 + \beta)^{-r}}{1 - (1 + \beta)^{-r}}$
Zero-Modified Logarithmic	β, p_0^M	$\beta^* = \theta\beta$, $p_0^{M*} = 1 - (1 - p_0^M) \frac{\ln(1 + \beta^*)}{\ln(1 + \beta)}$

Table 3: Parameter Transformation for Loss Count Distribution Fitting

4.3 Data Fitting Procedure

Suppose we have an arbitrary probabilistic model for raw loss size L and a parametric probabilistic model for raw loss count N^L which is assumed to satisfy theorem 4.1. Denote the distribution parameters as $\boldsymbol{\alpha} \in \mathbb{R}^m$. Given policy deductible d, we can compute $\theta = \mathbb{P}(L > d)$. If we observe data for coverage modified loss counts N^P of sample size n, denoted by $\mathbf{N^P} = \{N^{P(i)}\}_{i=1}^n$, then we can fit such data on our parametric model for N^P under the same parametric family using any acceptable method such as *Maximum Likelihood Estimation* (MLE), *Moments Matching Estimation* (MME), and *Percentile Matching Estimation* (PME). Denote N^P's parameter estimate as $\widehat{\boldsymbol{\alpha}}^*$. Finally, we transform the estimator back to obtain an estimator $\widehat{\boldsymbol{\alpha}}$ for N^L's parameters using theorem 4.1.

Figure 5: Data Fitting Procedure for Coverage Modified Frequency Random Variable

4.4 Numerical Example

In this section, we simulate data with the same specifications as section 3.10: sample size $n = 100$, ordinary deductible $d = \$5,000$, policy limit $u = \$20,000$, coinsurance factor $\alpha = 0.9$, under a constant inflation rate $r = 5\%$, with a log-normal raw loss size $\ln L \sim \mathcal{N}(\mu = 9, \sigma^2 = 1)$, and a zero-modified Poisson raw loss count $N^L \sim$ ZM-Poisson$(p_0^M = 0, \lambda = 5)$. We plot simulated count data below, for both raw and modified cases, with their respective theoretical PMFs.

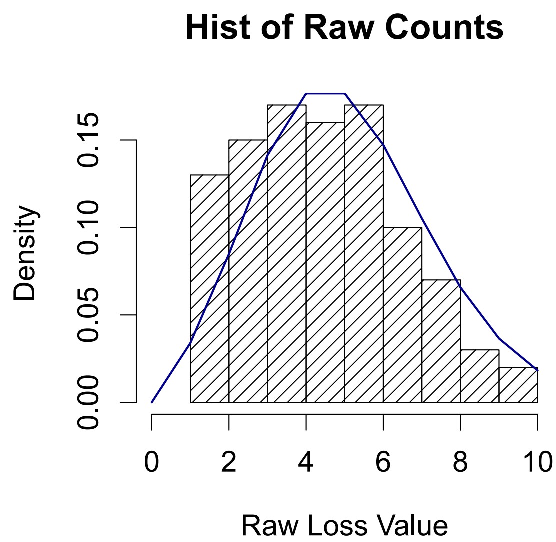


Figure 6: Example 4.4 Raw Loss Count

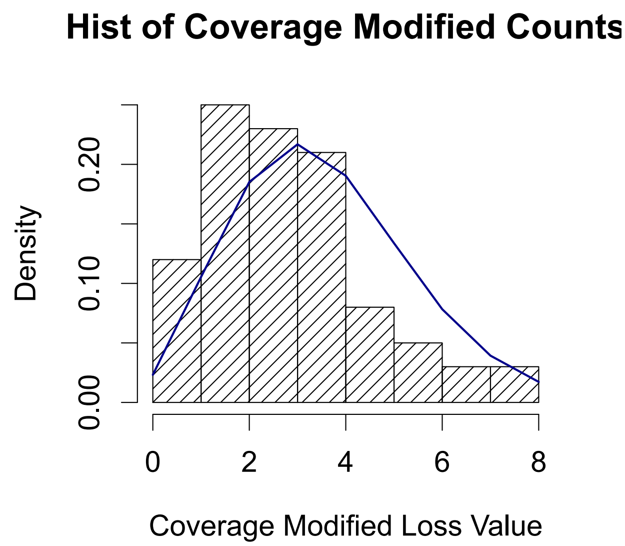


Figure 7: Example 4.4 Modified Loss Count

Given the above specification and simulated data, we can estimate parameters (p_0^M, λ) using MLE, MME with first and second moments, and PME with 33% and 66% percentiles. In addition, we perform bootstrap for $B = 1000$ times for each method to estimate the standard errors of the estimations. A summary is listed below.

	Maximum Likelihood			Moments Matching			Quantile Matching		
	Median	2.5%	97.5%	Median	2.5%	97.5%	Median	2.5%	97.5%
λ	4.6348	4.1008	5.1663	4.5371	3.9324	5.2128	7.2954	7.2954	7.5234
p_0^M	0.0112	-0.0315	0.0603	-0.0101	-0.1000	0.0724	0.2963	0.2319	0.2968

Table 4: Example 4.4 Estimated Parameters and Non-Parametric Bootstrap Confidence Intervals

Recall that the true parameter values are $(\lambda = 5, p_0^M = 0)$. We can observe from Table 4 that MLE is the most accurate estimate while QME is the least accurate. Notice that there are some estimated values for p_0^M that are outside its domain $[0, 1]$. (In particular, negative in this case.) This is due to the backward transformations ($\lambda^* \to \lambda$ and $p_0^{M*} \to p_0^M$) we need to apply at the end.

5 Concluding Remarks

In this paper, we presented derivations and results for probabilistic models of coverage modified severity and frequency random variables under all possible policy setups. As alluded to earlier, such theoretical formulation establishes a framework for future software development in insurance claim data fitting. One potential challenge, however, is to properly deal with numeric algorithms and starting points as they may cause failures in practical applications. Nonetheless, this paper is a comprehensive yet concise reference for researchers, practitioners, and software developers.

References

[Dutang et al., 2008] Dutang, C., Goulet, V., and Pigeon, M. (2008). actuar: An R Package for Actuarial Science. *Journal of Statistical Software*, 25(7):38.

[Klugman et al., 2012] Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2012). *Loss Models: From Data to Decisions*. John Wiley & Sons, Inc., 3rd edition.