Chapter 3: Frequency Models

Lecturer: Kenneth Ng

Preview

In this chapter, we are going to consider different models for claim frequency. We start with some basic discrete frequency distributions. Then, we shall look at some general classes of distributions whose pmfs satisfy some recursive relations – the $(a, b, 0)$ and $(a, b, 1)$ classes. We will further study more sophisticated compound models, and how deductibles would affect the payment frequency.

Key topics in this chapter:

1. Common parametric distributions for modelling frequency – Binomial, Negative Binomial, Poisson;
2. Mixed Poisson distribution;
3. The $(a, b, 0)$ class;
4. The $(a, b, 1)$ class – zero truncated and zero modified distributions;
5. Compound frequency model;
6. Effect of deductibles on payment frequency.

1 Basic Frequency Distributions

In what follows, we will let N be a *frequency variable*, i.e., a discrete random variable that describes the number of claims received by an insurer, whose support is given by $\mathbb{N}_0 := \{0, 1, 2, \dots\}$. In this section, we shall look at some common discrete, parametric distributions for modelling frequency.

1.1 Binomial Distribution

N follows a binomial distribution with parameters $m \in \mathbb{N}$ and $q \in [0, 1]$, denoted by $N \sim \text{Bin}(m, q)$, if it has the following pmf:

$$\boxed{p_N(k) = \binom{m}{q} q^k (1 - q)^{m-k},\ k = 0, 1, \dots, m.}$$

- The binomial distribution models the number of successes (e.g. number of claims) over m independent trials when the success rate (e.g. claim rate) of each trial is q.
- Some distributional quantities of $N \sim \text{Bin}(m, q)$:

> 1. **Mean and Variance:**
>
> $$\mathbb{E}[N] = mq, \quad \text{Var}[N] = mq(1-q).$$
>
> 2. **Probability Generating Function:**
>
> $$P_N(t) = (1 + q(t-1))^m, \ t > 0.$$

- The pgf can be derived as follows:

$$\begin{aligned} P_N(t) = \mathbb{E}[t^N] &= \sum_{k=0}^{m} \binom{m}{k} t^k q^k (1-q)^{m-k} = \sum_{k=0}^{m} \binom{m}{k} (tq)^k (1-q)^{m-k} \\ &= (tq + (1-q))^m = (1 + q(t-1))^m. \end{aligned}$$

- When $mq(1-q)$ is large (i.e. when m is large and q is away from 0 and 1), we can use a normal distribution to approximate a binomial distribution.
- $N \sim \text{Bin}(m, q)$ is the ***m*-convolution** (i.e. m independent sum) of Bernoulli distribution with parameter q:

> **Theorem 1.1** Let $N_1, \ldots, N_m$ be independent Bernoulli random variables with parameter q (i.e. $N_i \sim \text{Bern}(q) = \text{Bin}(1, q)$). Then, $N := \sum_{i=1}^{m} N_i \sim \text{Bin}(m, q)$.

Proof. We know that for each $i = 1, \ldots, m$, $P_{N_i}(t) = 1 + q(t-1)$. By independence, the pgf of N is given by

$$P_N(t) = \mathbb{E}\left[t^{N_1 + \cdots + N_m}\right] = \prod_{i=1}^{m} \mathbb{E}[t^{N_i}] = \prod_{i=1}^{m} P_{N_i}(t) = (1 + q(t-1))^m,$$

which is the pgf of $\text{Bin}(m, q)$. □

1.2 Negative Binomial Distribution

N follows a negative binomial distribution with parameters $r > 0$ and $\beta > 0$, denoted by $N \sim \text{NB}(r, \beta)$, if it has the following pmf:

$$\boxed{p_N(k) = \binom{r+k-1}{k} \left(\frac{1}{1+\beta}\right)^r \left(\frac{\beta}{1+\beta}\right)^k, \ k \in \mathbb{N}_0.}$$

Here, the binomial coefficient for $x \in \mathbb{R}$ and $k \in \mathbb{N}$ is defined as

$$\boxed{\binom{x}{k} = \frac{x(x-1)\cdots(x-k+1)}{k!}, \quad \binom{x}{0} = 1.}$$

When $x > k-1$, we can also write the binomial coefficient in terms of the gamma function:

$$\boxed{\binom{x}{k} = \frac{\Gamma(x+1)}{\Gamma(x-k+1)\Gamma(k+1)}.}$$

- When r is a positive integer, the negative binomial distribution models the number of failures in a sequence of Bernoulli trials until the r-th success occurs, where the failure odd is $\beta > 0$.
- Some distributional quantities of $N \sim \mathrm{NB}(r, \beta)$:

1. **Mean and Variance:**

$$\mathbb{E}[N] = r\beta, \quad \mathrm{Var}[N] = r\beta(1+\beta).$$

2. **Probability Generating Function:**

$$P_N(t) = (1-\beta(t-1))^{-r}, \; |t| < 1 + \frac{1}{\beta}.$$

- To show that p_N is a valid pmf, and to derive the pgf, we can use the following relation: for $\alpha \in \mathbb{R}$ and $|x| < 1$,

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \quad \text{and} \quad \binom{-\alpha}{k} = (-1)^k \binom{\alpha+k-1}{k}.$$

For instance, using this, the pgf can be shown as follows:

$$\begin{aligned}
P_N(t) = \mathbb{E}[t^N] &= \sum_{k=0}^{\infty} t^k \binom{r+k-1}{k} \left(\frac{1}{1+\beta}\right)^r \left(\frac{\beta}{1+\beta}\right)^k \\
&= \left(\frac{1}{1+\beta}\right)^r \sum_{k=0}^{\infty} \binom{r+k-1}{k} \left(\frac{t\beta}{1+\beta}\right)^k \\
&= \left(\frac{1}{1+\beta}\right)^r \sum_{k=0}^{\infty} \binom{-r}{k} \left(-\frac{t\beta}{1+\beta}\right)^k \\
&= \left(\frac{1}{1+\beta}\right)^r \left(1 - \frac{t\beta}{1+\beta}\right)^{-r} \quad \text{(for } |t| < 1 + 1/\beta) \\
&= (1-\beta(t-1))^{-r}.
\end{aligned}$$

- The ***geometric distribution*** is a special case of the negative binomial distribution when $r = 1$, which is denoted by $N \sim \text{Geom}(\beta) = \text{NB}(1, \beta)$.

1.3 Poisson Distribution

N follows a Poisson distribution with parameter $\lambda > 0$, denoted by $N \sim \text{Poi}(\lambda)$, if it has the following pmf:

$$\boxed{p_N(k) = \frac{e^{-\lambda}\lambda^k}{k!}, \; k \in \mathbb{N}_0.}$$

- Poisson distributions are used to model occurrences of events (e.g. insurance claims) with a constant rate of occurrences in a certain time period.
- It is the asymptotic limit of the binomial distribution when $m \to \infty$ and $mq \to \lambda$.
- Some distributional quantities of $N \sim \text{Poi}(\lambda)$:

1. **Mean and Variance:**
$$\mathbb{E}[N] = \lambda, \quad \text{Var}[N] = \lambda.$$
2. **Probability Generating Function:**
$$P_N(t) = e^{\lambda(t-1)}, t \in \mathbb{R}.$$

- The pgf can be derived as follows:

$$P_N(t) = \mathbb{E}[t^N] = \sum_{k=0}^{\infty} t^k \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(t\lambda)^k}{k!} = e^{-\lambda} e^{\lambda t} = e^{\lambda(t-1)}.$$

- $N \sim \text{Poi}(\lambda)$ is the ***m-convolution*** of Poisson distributions with parameters λ_i which sums to λ:

Theorem 1.2 Let N_i, $i = 1, \ldots, m$, follows $N_i \sim \text{Poi}(\lambda_i)$. Then, $N := \sum_{i=1}^n N_i \sim \text{Poi}(\sum_{i=1}^n \lambda_i)$.

Proof. For each $i = 1, 2, \ldots, m$, we know that the pgf of N_i $\text{Poi}(\lambda_i)$ is given by $P_{N_i}(t) = e^{\lambda_i(t-1)}$. By independence,

$$P_N(t) = \mathbb{E}\left[t^{N_1+\cdots+N_m}\right] = \prod_{i=1}^{m} \mathbb{E}[t^{N_i}] = \prod_{i=1}^{m} e^{\lambda_i(t-1)} = e^{(t-1)\sum_{i=1}^n \lambda_i},$$

which is the pgf of $\text{Poi}(\sum_{i=1}^n \lambda_i)$. □

- Conversely, if the number of claims N follows a Poisson distribution, and each claim can be classified into one of m different classes. Then, the number of claims for each class also follows a Poisson distribution. This observation is called ***thinning***. The precise statement and its proof are deferred to Theorem 6.2 below.

2 Mixed Poisson Distribution

In Chapter 2, we introduced mixture distributions. For instance, when the parameter of a distribution is random, the unconditional distribution can be viewed as a mixture of the conditional ones. In frequency model, if N is a random variable such that

$$\boxed{N|\Lambda = \lambda \sim \mathrm{Poi}(g(\lambda)),}$$

where Λ is also a random variable, we say that N has a ***mixed Poisson distribution***.

In a Poisson distribution, the mean and variance are equal, which can be limiting for modeling frequency data, as this assumption often does not hold in practice, where the variance is typically greater than the mean. In contrast, if N has a mixed Poisson distribution, we always have $\mathrm{Var}[N] > \mathbb{E}[N]$.

Proposition 2.1 Let N be a mixed Poisson variable, i.e., $N|\Lambda = \lambda \sim \mathrm{Poi}(g(\lambda))$. Then,

$$\boxed{\mathbb{E}[N] = \mathbb{E}[g(\Lambda)] \quad \text{and} \quad \mathrm{Var}[N] = \mathbb{E}[g(\Lambda)] + \mathrm{Var}[g(\Lambda)].}$$

In particular $\mathrm{Var}[N] > \mathbb{E}[N]$.

Proof. Since $N|\Lambda = \lambda \sim \mathrm{Poi}(g(\lambda))$, we have $\mathbb{E}[N|\Lambda = \lambda] = g(\lambda) = \mathrm{Var}[N|\Lambda = \lambda]$. By the law of iterated expectation,

$$\mathbb{E}[N] = \mathbb{E}[\mathbb{E}[N|\Lambda]] = \mathbb{E}[g(\Lambda)].$$

Next, by the law of total variance,

$$\mathrm{Var}[N] = \mathbb{E}[\mathrm{Var}[N|\Lambda]] + \mathrm{Var}[\mathbb{E}[N|\lambda]] = \mathbb{E}[g(\Lambda)] + \mathrm{Var}[g(\Lambda)].$$

□

2.1 Poisson-Gamma Mixture

One important case of a mixed Poisson distribution is when the mixing distribution Λ follows a gamma distribution, $\Lambda \sim \mathrm{Gamma}(\alpha, \theta)$. In this case, the mixed Poisson variable indeed follows a negative binomial distribution.

Theorem 2.2 Suppose that N is a mixed Poisson variable with $N|\Lambda \sim \mathrm{Poi}(\Lambda)$. If $\Lambda \sim \mathrm{Gamma}(\alpha, \theta)$, then $N \sim \mathrm{NB}(\alpha, \theta)$.

Proof. The pdf of $\Lambda \sim \text{Gamma}(\alpha, \theta)$ is given by

$$f_\Lambda(\lambda) = \frac{1}{\theta^\alpha \Gamma(\alpha)} \lambda^{\alpha-1} e^{-\frac{\lambda}{\theta}}, \ \lambda > 0.$$

On the other hand, for $k \in \mathbb{N}_0$,

$$\mathbb{P}(N = k | \Lambda = \lambda) = \frac{e^{-\lambda}\lambda^k}{k!}.$$

Using the law of total probability, we have, for any $k \in \mathbb{N}_0$,

$$\begin{aligned}
\mathbb{P}(N = k) &= \int_0^\infty \mathbb{P}(N = k|\Lambda = \lambda) f_\Lambda(\lambda) d\lambda \\
&= \int_0^\infty \left(\frac{e^{-\lambda}\lambda^k}{k!}\right) \left(\frac{1}{\theta^\alpha \Gamma(\alpha)} \lambda^{\alpha-1} e^{-\frac{\lambda}{\theta}}\right) d\lambda \\
&= \frac{1}{\theta^\alpha \Gamma(\alpha) k!} \int_0^\infty \lambda^{\alpha+k-1} e^{-\left(\frac{1+\theta}{\theta}\right)\lambda} d\lambda \\
&= \frac{1}{\theta^\alpha \Gamma(\alpha)} \left(\frac{\theta}{1+\theta}\right)^{\alpha+k} \int_0^\infty t^{\alpha+k-1} e^{-t} dt \\
&= \frac{\Gamma(\alpha+k)}{k!\Gamma(\alpha)} \left(\frac{1}{1+\theta}\right)^\alpha \left(\frac{\theta}{1+\theta}\right)^k \\
&= \binom{\alpha+k-1}{k} \left(\frac{\theta}{1+\theta}\right)^k \left(\frac{1}{1+\theta}\right)^\alpha,
\end{aligned}$$

where the fourth line follows from a change of variable $t = (1+\theta)\lambda/\theta$. From the above, we see that the pmf of N agrees with the pmf of $\text{NB}(\alpha, \theta)$. □

Example 2.1 The annual number of insurance claims has a Poisson distribution with mean Λ. The Poisson mean Λ follows a gamma distribution with mean 20 and variance 200. Find the probability that there are at least two claims in a year.

Solution:

It is known that $\Lambda \sim \text{Gamma}(\alpha, \theta)$, and

$$\mathbb{E}[\Lambda] = \alpha\theta = 20 \quad \text{and} \quad \text{Var}[\Lambda] = \alpha\theta^2 = 200.$$

Solving yields $\alpha = 2$ and $\theta = 10$. Hence, the number of claims $N \sim \text{NB}(2, 10)$, and

$$\begin{aligned}
\mathbb{P}(N \geq 2) &= 1 - \mathbb{P}(N = 0) - \mathbb{P}(N = 1) \\
&= 1 - \left(\frac{1}{1+10}\right)^2 - \binom{2}{1} \left(\frac{10}{1+10}\right) \left(\frac{1}{1+10}\right)^2 \\
&= 0.97671.
\end{aligned}$$

3 The $(a, b, 0)$ Class

As we will see, the pmfs of binomial, negative binomial, and the Poisson distributions introduced in the last section satisfy a special recursive relation. These distributions are said to belong to the $(a, b, 0)$ class, defined as follows:

Definition 3.1 Let $p_k = p_N(k)$, $k \in \mathbb{N}_0$ be the pmf of the random variable N. The distribution of N is a member of the $\boldsymbol{(a, b, 0)}$ ***class*** if there exist constants a and b such that

$$\boxed{\frac{p_k}{p_{k-1}} = a + \frac{b}{k}, \ k = 1, 2, \ldots} \tag{1}$$

The values a, b in the name $(a, b, 0)$ specify the constants in the recursive relation, and 0 indicates that the recursion starts at p_0. It is not difficult to verify that the binomial, negative binomial, and the Poisson distributions belong to the $(a, b, 0)$ class. The following table summarizes the constants a and b for these distributions:

1. **Binomial:**
$$a = -\frac{q}{1-q}, \ b = \frac{(m+1)q}{1-q},$$
2. **Negative Binomial:**
$$a = \frac{\beta}{1+\beta}, \ b = \frac{(r-1)\beta}{1+\beta},$$
3. **Poisson:**
$$a = 0, \ b = \lambda.$$

Theorem 3.1 The Poisson, negative binomial, and binomial distributions are the only members of the $(a, b, 0)$ class.

Theorem 3.1 can be proven by identifying the general form of the pgf of an $(a, b, 0)$ distribution, the details are out of the scope of our course.

Theorem 3.1, along with the table above it, allow us to identify a distribution from the $(a, b, 0)$ class:

Consider a distribution belonging to the $(a, b, 0)$ class:
- If $a > 0$, the distribution is negative binomial;
- If $a = 0$, the distribution is Poisson;
- If $a < 0$, the distribution is binomial.

Example 3.1 Suppose that the distribution of N belongs to the $(a, b, 0)$ class, with $p_0 = p_1 = 0.25$, and $p_2 = 0.1875$. Find p_3 and determine the distribution of N.

Solution:

Using the definition of $(a, b, 0)$ class, we know that

$$\begin{cases} \dfrac{0.25}{0.25} = \dfrac{p_1}{p_0} = a + \dfrac{b}{1} = a + b, \\ \dfrac{0.1875}{0.25} = \dfrac{p_2}{p_1} = a + \dfrac{b}{2} \end{cases} \Rightarrow \begin{cases} a + b = 1, \\ a + \dfrac{b}{2} = 0.75. \end{cases}$$

Solving yields $a = b = 0.5$. Hence,

$$p_3 = p_2 \left(0.5 + \frac{0.5}{3}\right) = 0.125.$$

Since $a = 0.5 > 0$, N follows a negative binomial distribution, where

$$0.5 = a = \frac{\beta}{1+\beta} \Rightarrow \beta = 1,$$

and

$$0.5 = b = \frac{(r-1)}{1+1} = \frac{r-1}{2} \Rightarrow r = 2,$$

In other words, $N \sim \text{NB}(2, 1)$.

Example 3.2 The pmf of a discrete distribution satisfies the following recursion:

$$p_k = \left(-c + \frac{4c}{k}\right) p_{k-1}, \; k = 1, 2, \ldots,$$

where $c \neq 0$. If $p_0 = 0.216$, find the value of c.

Solution:

Since $c \neq 0$, the distribution is either negative binomial or binomial. With $a = -c$ and $b = 4c$, $b/a = -4$. If the distribution is negative binomial, $b/a = r - 1 = -4$ if $r = -3$, which contradicts with the requirement that $r > 0$. Hence, the pmf is given by a binomial distribution. For a binomial distribution,

$$\frac{b}{a} = -(m+1) = -4,$$

which indicates that $m = 3$. By considering $0.216 = p_0 = (1-q)^3$, we have $q = 0.4$.

Therefore,

$$a = -c = -\frac{q}{1-q} = -\frac{0.4}{1-0.4} = -\frac{2}{3} \Rightarrow c = \frac{2}{3}.$$

4 The $(a, b, 1)$ Class

The $(a, b, 0)$ class provides a convenient characterization of some distributions for frequency models. Only, it also suffers from a few shortcomings:

- $(a, b, 0)$ class is too narrow and restrictive: it only consists of 3 distributions;
- For insurance claims, the probability of receiving 0 claim ($p_0 = \mathbb{P}(N = 0)$) is often very large. This could not be captured by the distributions in the $(a, b, 0)$ class.

To address this, we introduce the $(a, b, 1)$ class, which offers a higher flexibility for modelling while preserving the recursive relation of the pmf.

Definition 4.1 The distribution of N is a member of the $\boldsymbol{(a, b, 1)}$ ***class*** if there exist constants a and b such that

$$\boxed{\frac{p_k}{p_{k-1}} = a + \frac{b}{k}, \ \ k = 2, 3, \ldots} \tag{2}$$

Compared with the $(a, b, 0)$ class, the recursive relation in Equation (2) starts at p_1, whence the name $(a, b, 1)$. The value p_0 can be set arbitrarily. Then, the values $p_k, k \geq 2$, are *determined* in a way that the $(a, b, 1)$ equation (2), and the following are satisfied:

$$\boxed{p_0 = 1 - \sum_{k=1}^{\infty} p_k.}$$

We consider two types of $(a, b, 1)$ class:

1. **Zero-Truncated** $(a, b, 1)$ **class**: The probability at 0 is set as $p_0 := 0$. The associated pmf is denoted by p_k^T, $k = 1, 2, \ldots$;
2. **Zero-Modified** $(a, b, 1)$ **class**: The probability at 0 is set as $p_0^M > 0$. The associated pmf is denoted by p_k^M, $k = 1, 2, \ldots$. Indeed, the zero-truncated class can also be considered as a special example of the zero-modified class.

4.1 Zero-Modified $(a, b, 1)$ Class

An $(a, b, 1)$ distribution can be constructed from an $(a, b, 0)$ distribution by the following:

Proposition 4.1 Let $\{p_k\}_{k=0}^{\infty}$ be a pmf which satisfies the $(a,b,0)$ equation (1). Define $\{p_k^M\}_{k=0}^{\infty}$ by

$$\boxed{\begin{aligned} p_k^M &:= cp_k, k=1,2,\dots,\\ c &:= \frac{1-p_0^M}{1-p_0}, \end{aligned}} \tag{3}$$

where $p_0^M \in [0,1)$ is arbitrarily chosen. Then, $\{p_k^M\}_{k=0}^{\infty}$ is a pmf satisfying the $(a,b,1)$ equation (2).

Proof. If $\{p_k^M\}_{k=0}^{\infty}$ satisfies Equation (3), then for any $k \geq 2$,

$$\frac{p_k^M}{p_{k-1}^M} = \frac{cp_k}{cp_{k-1}} = \frac{p_k}{p_{k-1}} = a + \frac{b}{k}.$$

In addition,

$$p_0^M + \sum_{k=1}^{\infty} p_k^M = p_0^M + \frac{1-p_0^M}{1-p_0} \sum_{k=1}^{\infty} p_k = p_0^M + \frac{1-p_0^M}{1-p_0}(1-p_0) = 1.$$

Therefore, $\{p_k^M\}_{k=0}^{\infty}$ is a pmf satisfying the $(a,b,1)$ equation. □

As a consequence of Proposition 4.1, an $(a,b,1)$ distribution can be constructed by following the steps below:

1. Identify a, b, and the associated $(a,b,0)$ distribution;
2. Set p_0^M and compute c;
3. Set $p_k^M := cp_k$.

Recall that the $(a,b,0)$ class only consists of three elements – binomial, negative binomial, and Poisson. The $(a,b,1)$ distribution constructed from these distributions are thus called ***zero-modified binomial*** (ZM-Bin(m,q,p_0^M)), ***zero-modified negative binomial*** (ZM-NB(r,β,p_0^M)), and ***zero-modified Poisson*** (ZM-Poi(λ,p_0^M)) distribution, respectively.

The pgfs of $(a,b,0)$ and $(a,b,1)$ distributions are related as follows:

Proposition 4.2 Let $P^M(t)$ be the pgf of an $(a,b,1)$ distribution, and $P(t)$ be the pgf of the associated $(a,b,0)$ distribution. Then,

$$\boxed{P^M(t) = 1 - c + cP(t).}$$

Proof. By Equation (3), we have

$$\begin{aligned}
P^M(t) &= p_0^M + \sum_{k=1}^{\infty} p_k^M t^k \\
&= p_0^M + c\sum_{k=1}^{\infty} p_k t^k \\
&= p_0^M + c\left(P(t) - p_0\right) \\
&= p_0^M - cp_0 + cP(t) \\
&= p_0^M - \frac{p_0(1-p_0^M)}{1-p_0} + cP(t) \\
&= \frac{p_0^M - p_0}{1-p_0} + cP(t) \\
&= 1 - c + cP(t).
\end{aligned}$$

□

As a consequence of Proposition 4.1, if we let μ and σ^2 be the mean and variance of an $(a, b, 0)$ distribution, the mean μ_M and variance σ_M^2 of the $(a, b, 1)$ distribution constructed from it are respectively given by

$$\boxed{\mu_M = c\mu \quad \text{and} \quad \sigma_M^2 = c(1-c)\mu^2 + c\sigma^2.}$$

4.2 Zero-Truncated $(a, b, 1)$ Class

The zero-truncated $(a, b, 1)$ class can be considered as a special case of the zero-modified class, with $p_0^T = 0$. Hence, by Equation (3), if $\{p_k\}_{k=1}^{\infty}$ is the pmf of an $(a, b, 0)$ distribution, then $\{p_k^T\}_{k=0}^{\infty}$ is the pmf of an $(a, b, 1)$ distribution if and only if the following holds:

$$\boxed{\begin{aligned} p_k^T &= cp_k, k = 1, 2, \ldots, \\ c &= \frac{1}{1-p_0}. \end{aligned}}$$

The pgf of a zero-truncated $(a, b, 1)$ distribution, $P^T(t)$, is given by

$$\boxed{P^T(t) = 1 - c + cP(t) = \left(1 - \frac{1}{1-p_0}\right) + \frac{P(t)}{1-p_0},}$$

where $P(t)$ is the pgf of the associated $(a, b, 0)$ distribution.

Example 4.1 The pmf of a discrete distribution satisfies the following recursion:

$$p_k^M = \frac{4p_{k-1}^M}{k}, \; k = 2, 3, \ldots.$$

If $p_0^M = 0.2$, find p_5^M.

<u>*Solution:*</u>

The pmf satisfies the $(a, b, 1)$ equation with $a = 0$ and $b = 4$. This indicates that the distribution is given by a zero-modified Poisson distribution with $\lambda = 4$. If we let $\{p_k\}_{k=0}^{\infty}$ be the pmf of Poi(4), we have $p_0 = e^{-4}$. Hence,

$$c = \frac{1 - p_0^M}{1 - p_0} = \frac{0.8}{1 - e^{-4}}.$$

The required probability is thus

$$p_5^M = cp_5 = \frac{0.8}{1 - e^{-4}} \left(\frac{e^{-4} 4^5}{5!} \right) = 0.1274.$$

Example 4.2 Suppose that the pgf of the random variable N is given by

$$P_N(t) = \frac{3}{4} + \frac{1}{4(3 - 2t)^3}, \; t < \frac{3}{2}.$$

Determine the distribution of N as a member of the $(a, b, 1)$ class.

<u>*Solution:*</u>

We can write

$$P_N(t) = \left(1 - \frac{1}{4}\right) + \frac{1}{4} \frac{1}{(3 - 2t)^3} = 1 - c + cP(t),$$

where $c = 1/4$, and $P(t) = (3 - 2t)^{-3}$ is the pgf of NB$(r = 3, \beta = 2)$. With $p_0 = P_N(0) = 1/27$, we can solve p_0^M by setting

$$\frac{1}{4} = c = \frac{1 - p_0^M}{1 - p_0} = \frac{1 - p_0^M}{1 - \frac{1}{27}} \Rightarrow p_0^M = \frac{41}{54}.$$

Therefore, $N \sim$ ZM-NB$(r = 3, \beta = 2, p_0^M = 41/54)$.

4.3 Extended Truncated Negative Binomial Distribution

Recall that the negative binomial distribution, $\mathrm{NB}(r,\beta)$, is parametrized by $r>0$ and $\beta>0$. If $r\in(-1,0)$,

$$p_0=\left(\frac{1}{1+\beta}\right)^r>1,\ p_1=r\left(\frac{1}{1+\beta}\right)^r\left(\frac{\beta}{1+\beta}\right)<0,$$

and for $k\geq 2$,

$$p_k=\frac{\beta}{1+\beta}\left(1+\frac{r-1}{k}\right)p_{k-1}<0,$$

which violates the requirement of being a pmf. However, it still holds that

$$\sum_{k=0}^{\infty}p_k=1.$$

The flexibility of setting $p_0^T=0$ allows us to extend the choice of parameter $r>0$, to $r>-1, r\neq 0$, that gives an $(a,b,1)$ distribution satisfying the same recursion relation (2) as a negative binomial distribution. The resulting distribution is referred to as the ***extended truncated negative binomial (ETNB) distribution***.

Definition 4.2 A random variable is said to follow an ***extended truncated negative binomial (ETNB) distribution***, denoted by $\mathrm{ETNB}(r,\beta)$, where $r>-1, r\neq 0$, and $\beta>0$, if its pmf $\{p_k^T\}_{k=0}^{\infty}$ is given by

$$\boxed{p_0^T=0,\ p_k^T=cp_k,\ k=1,2,\ldots,} \tag{4}$$

where

$$\boxed{\begin{aligned} c&:=\frac{1}{1-\left(\frac{1}{1+\beta}\right)^r},\\ p_k&:=\binom{r+k-1}{k}\left(\frac{1}{1+\beta}\right)^r\left(\frac{\beta}{1+\beta}\right)^k,\ k=1,2,\ldots \end{aligned}} \tag{5}$$

To see that Equations (4) - (5) indeed define a valid pmf, consider

$$\sum_{k=0}^{\infty}p_k^T=c\sum_{k=1}^{\infty}p_k=c\left(1-\left(\frac{1}{1+\beta}\right)^r\right)=1.$$

If $r>0$, we have $c>0$ and $p_k>0$, and thus $p_k^T>0$ for any $k\in\mathbb{N}$. If $-1<r<0$, we have $c<0$ and $p_k<0$, whence we still have $p_k^T>0$. Therefore, $\{p_k^T\}_{k=1}^{\infty}$ is indeed a valid pmf.

5 Compound Frequency Models

A *compound model* is given by the sum of N random variables, where N itself is also random. In frequency models, we can interpret N as the number of accidents within a period of time, and M_k, $k = 1, 2, \ldots, N$, be the random variable of the number of claims received associated to the k-th accident. Hence, we can model the total number of claims S by:

$$\boxed{S := M_1 + M_2 + \cdots + M_N = \sum_{k=1}^{N} M_k.}$$

The random variable S is said to follow a ***compound distribution***. In this chapter, we assume $M_1, M_2, \ldots$ are i.i.d. random variables with a common distribution represented by M, and are independent of N. For the compound variable S, we call the distribution of N as the ***primary distribution***, and the distribution of M as the ***secondary distribution***.

It is not an easy task to express the pmf of S in a neat manner. Alternatively, we can characterize its distribution of the pgf of M and N:

Theorem 5.1 The pgf of S is given by

$$\boxed{P_S(t) = P_N(P_M(t)).}$$

Proof. By the definition of $P_S(t)$, the independence of M and N, and the law of iterated expectation, we have

$$\begin{aligned}
P_S(t) &= \mathbb{E}[t^S] = \mathbb{E}\left[\mathbb{E}[t^S|N]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[t^{M_1+\cdots+M_N}|N\right]\right] \\
&= \mathbb{E}\left[\prod_{k=1}^{N} \mathbb{E}[t^{M_k}|N]\right] \\
&= \mathbb{E}\left[(\mathbb{E}[t^M])^N\right] \\
&= \mathbb{E}\left[(P_M(t))^N\right] \\
&= P_N(P_M(t)).
\end{aligned}$$

Notice that the third-to-last line follows since $M_1, \ldots, M_n$ are i.i.d. with a common distribution M, and are independent of N. □

Using the pgf of S, or alternatively, by the law of iterated expectation/law of total variance, the mean and variance of S is given as follows:

Proposition 5.2 For the compound variable S, we have

$$\boxed{\mathbb{E}[S] = \mathbb{E}[N]\mathbb{E}[M] \quad \text{and} \quad \mathrm{Var}[S] = \mathbb{E}[N]\mathrm{Var}[M] + \mathbb{E}^2[M]\mathrm{Var}[N].}$$

Proof. To compute $\mathbb{E}[S]$, we apply the law of iterated expectation:

$$\mathbb{E}[S] = \mathbb{E}[\mathbb{E}[S|N]] = \mathbb{E}\left[\mathbb{E}[M_1 + \cdots + M_N|N]\right] = \mathbb{E}[N\mathbb{E}[M]] = \mathbb{E}[M]\mathbb{E}[N],$$

where the second last equality follows from the facts that $M \perp\!\!\!\perp N$, and M_1, M_2 are i.i.d. To find $\text{Var}[S]$, we apply the law of total variance, and make use of the independence of M and N to obtain

$$\begin{aligned}\text{Var}[S] &= \mathbb{E}[\text{Var}[S|N]] + \text{Var}[\mathbb{E}[S|N]]\\ &= \mathbb{E}[\text{Var}[M_1 + \cdots + M_N|N]] + \text{Var}[\mathbb{E}[M_1 + \cdots + M_N|N]]\\ &= \mathbb{E}[N\text{Var}[M]] + \text{Var}[N\mathbb{E}[M]]\\ &= \mathbb{E}[N]\text{Var}[M] + \mathbb{E}^2[M]\text{Var}[N].\end{aligned}$$

□

Example 5.1 Let S be a compound random variable with the following pgf:

$$P_S(t) = \exp\left(\frac{5}{(1-3(t-1))^4} - 5\right), \quad |t| < \frac{4}{3}.$$

(a) Identify the primary and secondary distribution of S.
(b) Find $\mathbb{E}[S]$ and $\text{Var}[S]$.

Solution:

(a) We can write $P_S(t)$ as

$$P_S(t) = e^{5(P_M(t)-1)} = P_N(P_M(t)),$$

where

$$P_N(t) = e^{5(t-1)} \quad \text{and} \quad P_M(t) = \frac{1}{(1-3(t-1))^4}.$$

Hence, we can identify the primary distribution is given by $N \sim \text{Poi}(5)$, and the secondary distribution is given by $M \sim \text{NB}(4, 3)$.

(b) To compute $\mathbb{E}[S]$, we have

$$\mathbb{E}[S] = \mathbb{E}[M]\mathbb{E}[N] = (4 \times 3)(5) = 60.$$

To compute $\text{Var}[S]$, notice that

$$\text{Var}[N] = 5 \quad \text{and} \quad \text{Var}[M] = 4 \times 3 \times (1+3) = 48.$$

Hence,

$$\text{Var}[S] = \mathbb{E}[N]\text{Var}[M] + \mathbb{E}^2[M]\text{Var}[N] = (5)(48) + (12)^2(5) = 960.$$

6 Payment Frequency under Deductibles

In the previous sections, we have discussed models for claim frequency N. When deductibles are imposed, not all claims will result in a payment. In this section, we use N^L to denote the *number of claims/loss*, and N^P to denote the *number of payments*.

Assume that each claim has a probability v of leading to a payment, e.g., $v = \mathbb{P}(X > d)$, where X is the loss size and d is the deductible. Then, we can write N^P as a compound distribution with N^L being the primary distribution:

$$\boxed{N^P = I_1 + I_2 + \cdots + I_N = \sum_{k=1}^{N} I_k,}$$

where for each $k = 1, 2, \ldots, N$, I_k is an indicator variable defined by

$$\boxed{I_k := \begin{cases} 1, & \text{if the } k\text{-th claim results in a payment;} \\ 0, & \text{otherwise.} \end{cases}}$$

Then, it is clear that $I_k \sim \text{Bernoulli}(v)$. Suppose that I_k, $k = 1, 2, \ldots,$ are independent, and are independent of N. Then we can express the pgf of N^P in terms of N^L and v:

Proposition 6.1 Suppose that I_k, $k = 1, 2, \ldots$ are independent, and are independent of N. Then, the pgf of N^P is given by

$$\boxed{P_{N^P}(t) = P_{N^L}(1 - v + vt).}$$

Proof. Let $I \sim \text{Bernoulli}(v) = \text{Bin}(1, v)$ be the representative variable of I_k, $k = 1, 2, \ldots$. By the pgf of a Bernoulli (or binomial) distribution, we know that

$$P_I(t) = 1 - v + vt.$$

By Theorem 5.1, we have

$$P_{N^P}(t) = P_{N^L}(P_I(t)) = P_{N^L}(1 - v + vt).$$

□

6.1 Special Case: $(a, b, 0)$ and $(a, b, 1)$ Classes

In general, it is not easy to find the distribution of N^P from N^L. However, if the distribution of N^L belongs to the $(a, b, 0)$ class, N^P will also belong to the same distribution family with different parameters. The following table summarizes the distribution of N^P given that of N^L, which can be easily proven by Proposition 6.1 and using the pgfs of the $(a, b, 0)$ distributions.

Distribution	N^L	N^P
Binomial	$\text{Bin}(m, q)$	$\text{Bin}(m, qv)$
Negative Binomial	$\text{NB}(r, \beta)$	$\text{NB}(r, \beta v)$
Poisson	$\text{Poi}(\lambda)$	$\text{Poi}(v\lambda)$

Table 1: Distribution of N^P if N^L belongs to the $(a, b, 0)$ class

Likewise, if N^L belongs to the $(a, b, 1)$ class, N^P will have the same type of distribution but with different parameters, and a different probability at zero. To see this, recall that the pgf of N^L is given by

$$P_{N^L}(t) = 1 - c + cP_N(t) = \left(1 - \frac{1 - p_0^M}{1 - p_0}\right) + \frac{1 - p_0^M}{1 - p_0} P_N(t),$$

where N belongs to the $(a, b, 0)$ distribution with $p_0 = \mathbb{P}(N = 0)$, and $p_0^M = \mathbb{P}(N^L = 0)$ is the modified zero probability of N^L. By Proposition 6.1, the pgf of N^P is given by

$$P_{N^P}(t) = P_{N^L}(1 - v + vt) = 1 - c + cP_N(1 - v + vt) = 1 - c + cP_{N^*}(t),$$

where N^* is the *revised* $(a, b, 0)$ distribution of N due to deductible, see Table 1. Thus, N^P belongs to the $(a, b, 1)$ class with the associated $(a, b, 0)$ variable N^*. The zero probability of N^P, $p_0^{M^*} := \mathbb{P}(N^P = 0)$, will also change due to the deductible revision, which can be computed by as follows:

$$p_0^{M*} = P_{N^P}(0) = 1 - c + cP_{N^*}(0) = 1 - c + cp_0^* = 1 - \frac{1 - p_0^M}{1 - p_0} + \frac{1 - p_0^M}{1 - p_0} p_0^*,$$

where $p_0^* := \mathbb{P}(N^* = 0)$. Equivalently,

$$\boxed{1 - p_0^{M^*} = \frac{1 - p_0^M}{1 - p_0}(1 - p_0^*).} \tag{6}$$

Table 2 below summarizes the relationship of N^L and N^P if N^L belongs to the $(a, b, 1)$ class from the above discussions.

Distribution	N^L	N^P
ZM-Binomial	ZM-Bin(m, q, p_0^M)	ZM-Bin(m, qv, p_0^{M*})
ZM-Negative Binomial	ZM-NB(r, β, p_0^M)	ZM-NB$(r, \beta v, p_0^{M*})$
ZM-Poisson	ZM-Poi(λ, p_0^M)	ZM-Poi$(v\lambda, p_0^{M*})$

	$(a, b, 1)$ variable	Associated $(a, b, 0)$ variable
No. of Loss	N^L, p_0^M	N, p_0
No. of Payment	N^P, p_0^{M*}	N^*, p_0^*

Table 2: Distribution of N^P if N^L belongs to the $(a, b, 1)$ class

Example 6.1 (**SOA EXAM STAM SAMPLE Q126 MODIFIED**) The number of annual losses has a Poisson distribution with a mean of 5. The size of each loss has a two-parameter Pareto distribution with $\theta = 10$ and $\alpha = 2.5$. An insurance for the losses has an ordinary deductible of 5 per loss. Calculate the expected number of payment.

Solution:

Since $X \sim \text{Pareto}(2.5, 10)$, the probability of payment is given by

$$v = \mathbb{P}(X > d) = \mathbb{P}(X > 5) = \left(\frac{10}{5 + 10}\right)^{2.5} = 0.36289.$$

Hence, the distribution of the number of payment N^P is given by $N^P = \text{Poi}(\lambda v) = \text{Poi}(5 \times 0.36289) = \text{Poi}(1.8144)$. Thus, the expected number of payment is 1.8144.

Example 6.2 The number of annual losses N^L has a zero-modified binomial distribution with $m = 10$, $q = 0.3$, and $\mathbb{P}(N^L = 0) = 0.6$. The size of each loss has an exponential distribution with mean 100. Suppose that a deductible of amount 70 is applied.

(a) Calculate the probability that there is no payment to be made.

(b) Calculate the expected number of payment.

Solution:

(a) First, we compute the probability of payment for each loss. For $X \sim \text{Exp}(100)$, we have

$$v = \mathbb{P}(X > 70) = e^{-\frac{70}{100}} = e^{-0.7}.$$

Hence, N^P follows a zero-modified binomial distribution with $m = 10$, and $q = 0.3e^{-0.7}$. Next, we calculate the modified zero probability p_0^{M*} of N^P. To this end, let $N^* \sim \text{Bin}(10, 0.3e^{-0.7})$, which is the associated $(a, b, 0)$ variable for N^P.

Consider

$$p_0 = \mathbb{P}(\text{Bin}(10, 0.3) = 0) = 0.7^{10},$$
$$p_0^* = \mathbb{P}(N^* = 0) = (1 - 0.3e^{-0.7})^{10}.$$

Hence,

$$1 - p_0^{M*} = \frac{1 - p_0^M}{1 - p_0}(1 - p_0^*) = \frac{1 - 0.6}{1 - 0.7^{10}}\left(1 - (1 - 0.3e^{-0.7})^{10}\right) = 0.32961,$$

which gives $p_0^{M*} = 1 - 0.32961 = 0.67039$.

(b) First, we have

$$\mathbb{E}[N^*] = 10(0.3e^{-0.7}) = 3e^{-0.7}.$$

Since N^* is the associated $(a, b, 0)$ variable for N^P, we can calculate the proportionality constant c of N^* and N^P by

$$c = \frac{1 - p_0^{M*}}{1 - p_0^*} = \frac{1 - p_0^M}{1 - p_0} = \frac{1 - 0.6}{1 - 0.7^{10}} = 0.41163.$$

Therefore,

$$\mathbb{E}[N^P] = c\mathbb{E}[N^*] = (0.41163)(3e^{-0.7}) = 0.61323.$$

6.2 General Thinning Theorems for $(a, b, 0)$ Class

Revising the loss frequency N^L to payment frequency N^P is an example of ***thinning***: each loss can be classified into either of the following:

1. loss that needs a payment, with probability $v = \mathbb{P}(X > d)$;
2. loss that no payment is needed, with probability $1 - v$.

This classification can be generalized to more than two types as follows.

Suppose that the number of claims is N, and each claim can be classified into one of m types with probability $p_1, \ldots, p_m$, where $\sum_{i=1}^m p_i = 1$. Let N_i, $i = 1, \ldots, m$ be the number of claims that belong to type i. Then, by following the proof of Proposition 6.1, the pgf of N_i is given by

$$\boxed{P_{N_i}(t) = P_N(1 - p_i + p_i t).}$$

As before, if N belongs to the $(a, b, 0)$ class, each N_i will have the same class of distribution, but with a different parameter as shown in Table 3 below.

In particular, if $N \sim \text{Poi}(\lambda)$, the *thinned variables* $N_1, \ldots, N_m$ will be *independent*; see Theorem 6.2 below. This is NOT true for binomial and negative binomial. Indeed, the

Distribution	N	N_i
Binomial	$\text{Bin}(m, q)$	$\text{Bin}(m, p_i v)$
Negative Binomial	$\text{NB}(r, \beta)$	$\text{NB}(r, \beta p_i)$
Poisson	$\text{Poi}(\lambda)$	$\text{Poi}(p_i \lambda)$

Table 3: Distribution of N_i, $i = 1, \ldots, m$, if N belongs to the $(a, b, 0)$ class

independence is a consequence of the superposition property of that is unique to the Poisson distribution (i.e., sum of independence Poisson is again a Poisson).

Theorem 6.2 (Thinning) Suppose that the number of claims $N \sim \text{Poi}(\lambda)$. Then, the thinned variables $N_1, \ldots, N_m$ are independent Poisson random variables with parameters $\lambda_1, \ldots, \lambda_m$, respectively.

Sketch of Proof. Using the fact that $P_{N_i}(t) = P_N(1 - p_i + p_i t)$, it can be shown easily that

$$P_{N_i}(t) = e^{\lambda p_i (t-1)},$$

which implies $N_i \sim \text{Poi}(\lambda p_i)$. To show that $N_1, \ldots, N_m$ are independent, it suffices to show that

$$P_{N_1 + \cdots + N_m}(t) = \prod_{i=1}^{m} P_{N_i}(t).$$

The factorization of the characteristic function implies that the distribution of $N_1 + \cdots + N_m$ is consistent with the distribution of their independent sum.

Indeed,

$$P_{N_1 + \cdots + N_m}(t) = P_N(t) = e^{\lambda(t-1)},$$

and

$$\prod_{i=1}^{m} P_{N_i}(t) = \prod_{i=1}^{m} e^{\lambda p_i (t-1)} = e^{\lambda(t-1) \sum_{i=1}^{m} p_i} = e^{\lambda(t-1)}.$$

Therefore, we conclude that $N_1, \ldots, N_m$ are independent[1].

□

[1]Strictly speaking, one needs to show that $P_{a_1 N_1 + \cdots + a_m N_m}(t)$ is the pgf of $\text{Poi}(a_1 \lambda_1 + \cdots + a_m \lambda_m)$ for all $(a_1, \ldots, a_n) \in \mathbb{R}^n$