

M2 Collective Risk Modelling

General Insurance Modelling : Actuarial Modelling III ¹

Professor Benjamin Avanzi



06 March 2023

- 1 Introduction: Models for aggregate losses
- 2 The Individual Risk Model
- 3 The Collective Risk Model (Compound distributions, MW 2.1)
- 4 Explicit claims count distributions (MW 2.2)
- 5 Additional properties and applications of Poisson frequencies
- 6 ✕ Parameter estimation (MW 2.3)
- 7 The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions

- 1 Introduction: Models for aggregate losses
- 2 The Individual Risk Model
- 3 The Collective Risk Model (Compound distributions, MW 2.1)
- 4 Explicit claims count distributions (MW 2.2)
- 5 Additional properties and applications of Poisson frequencies
- 6 ✕ Parameter estimation (MW 2.3)
- 7 The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions

Introduction: Models for aggregate losses

A portfolio of contracts or a contract will potentially experience a sequence of losses:

$$Y_1, Y_2, Y_3, \dots$$

We are interested in the aggregate sum S of these losses over a certain period of time.

- How many losses will occur?
 - if deterministic (n) \rightarrow individual risk model
 - if random (N) \rightarrow collective risk model
- How do they relate to each other?
 - usual assumption: iid
- When do these losses occur?
 - usual assumption: no time value of money
 \rightarrow short term models
- How big are these losses?

- 1 Introduction: Models for aggregate losses
- 2 The Individual Risk Model**
- 3 The Collective Risk Model (Compound distributions, MW 2.1)
- 4 Explicit claims count distributions (MW 2.2)
- 5 Additional properties and applications of Poisson frequencies
- 6 ✕ Parameter estimation (MW 2.3)
- 7 The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions

2 The Individual Risk Model

- Definition
- Convolutions of random variables
- Using generating functions
- Using R

The Individual Risk Model

In the Individual Risk Model

$$S = Y_1 + \cdots + Y_n = \sum_{i=1}^n Y_i,$$

where Y_i , $i = 1, 2, \dots, n$, are iid claims. There are several methods to get probabilities about S :

- get the whole distribution of S (if possible)
 - Convolutions
 - Generating functions
- (⊠) approximate with the help of the moments of S (Module 4)

2 The Individual Risk Model

- Definition
- Convolutions of random variables
- Using generating functions
- Using R

Convolutions of random variables

In probability, the operation of determining the distribution of the sum of two random variables is called a **convolution**. It is denoted by

$$F_{X+Y} = F_X * F_Y.$$

The result can then be convolved with the distribution of another random variable. For instance,

$$F_{X+Y+Z} = F_Z * F_{X+Y}.$$

This can be done for both discrete and continuous random variables. It is also possible for mixed rv's, but it is more complicated.

Formulas

In short

- Discrete case:
 - df: $F_{X+Y}(s) = \sum_x F_Y(s-x) f_X(x)$
 - pmf: $f_{X+Y}(s) = \sum_x f_Y(s-x) f_X(x)$
- Continuous case:
 - cdf: $F_{X+Y}(s) = \int_{-\infty}^s F_Y(s-x) f_X(x) dx$
 - pdf: $f_{X+Y}(s) = \int_{-\infty}^s f_Y(s-x) f_X(x) dx$

Examples:

- discrete case: Bowers et al. (1997) Example 2.3.1 on page 35
- continuous case: Bowers et al. (1997) Example 2.3.2 on page 36

Numerical example

Consider 3 discrete r.v.'s with probability mass functions

$$\begin{aligned}f_1(y) &= \frac{1}{4}, \frac{1}{2}, \frac{1}{4} && \text{for } y = 0, 1, 2 \\f_2(y) &= \frac{1}{2}, \frac{1}{2} && \text{for } y = 0, 2 \\f_3(y) &= \frac{1}{4}, \frac{1}{2}, \frac{1}{4} && \text{for } y = 0, 2, 4\end{aligned}$$

Calculate the pmf f_{1+2+3} and the df F_{1+2+3} of the sum of the three random variables.

Solution

y	$f_1(y)$	$f_2(y)$	$f_{1+2}(y)$	$f_3(y)$	$f_{1+2+3}(y)$	$F_{1+2+3}(y)$
0	1/4	1/2	1/8	1/4	1/32	1/32
1	1/2	0	2/8	0	2/32	3/32
2	1/4	1/2	2/8	1/2	4/32	7/32
3	0	0	2/8	0	6/32	13/32
4	0	0	1/8	1/4	6/32	19/32
5	0	0	0	0	6/32	25/32
6	0	0	0	0	4/32	29/32
7	0	0	0	0	2/32	31/32
8	0	0	0	0	1/32	32/32

$$f_{1+2}(2) = 1/4 \cdot 1/2 + 1/2 \cdot 0 + 1/4 \cdot 1/2$$

$$f_{1+2+3}(4) = 1/8 \cdot 1/4 + 2/8 \cdot 0 + 2/8 \cdot 1/2 + 2/8 \cdot 0 + 1/8 \cdot 1/4$$

2 The Individual Risk Model

- Definition
- Convolutions of random variables
- Using generating functions
- Using R

Using generating functions

There is a 1-1 relation between a distribution and its mgf or pgf.

Because

$$M_S(t) = E[e^{tS}] = E[e^{t(Y_1 + \dots + Y_n)}] = E[e^{tY_1} \dots e^{tY_n}]$$

and if losses are independent then we have

$$M_S(t) = E[e^{tS}] = E[e^{tY_1}] \dots E[e^{tY_n}] = M_{Y_1}(t) \dots M_{Y_n}(t).$$

The same argument holds for the pgf's.

- Sometimes, $M_S(t)$ or $p_S(t)$ can be recognised: this is the case for infinitely divisible distributions (Normal, Poisson, Inverse Gaussian, ...) and certain other distributions (Binomial, Negative binomial).
- Otherwise, $M_S(t)$ or $p_S(t)$ can be expanded numerically to get moments and/or probabilities.

Example

Consider a portfolio of 10 contracts. The losses Y_i 's for these contracts are iid rv's with mean 100 and variance 100. Determine the distribution, the expected value and the variance of S if these losses are

- ① Normal;
- ② Gamma;
- ③ Poisson.

2 The Individual Risk Model

- Definition
- Convolutions of random variables
- Using generating functions
- Using R

Using R

- Contrary to Excel, convolutions are extremely easy to implement in R using vectors.

```
f1 <- c(1/4, 1/2, 1/4, 0, 0)
f2 <- c(1/2, 0, 1/2, 0, 0)
f12 <- c(f1[1] * f2[1], sum(f1[1:2] * f2[2:1]), sum(f1[1:3] *
  f2[3:1]), sum(f1[1:4] * f2[4:1]), sum(f1[1:5] * f2[5:1]))
f12
```

```
## [1] 0.125 0.250 0.250 0.250 0.125
```

- The example above is generalised in Exercise 1os9R.
- A more advanced R function is `convolve`. It actually involves the Fast Fourier Transform (a method that is related to that of the mgf's) for efficiency. We do not discuss this here, but it is used in the implementation of convolutions in the function `aggregateDist` of the package `actuar` (introduced later).

- 1 Introduction: Models for aggregate losses
- 2 The Individual Risk Model
- 3 The Collective Risk Model (Compound distributions, MW 2.1)**
- 4 Explicit claims count distributions (MW 2.2)
- 5 Additional properties and applications of Poisson frequencies
- 6 ✕ Parameter estimation (MW 2.3)
- 7 The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions

3 The Collective Risk Model (Compound distributions, MW 2.1)

- Definition
- Moments of S
- Moment generating function of S
- Distribution of S
- Using R

Introduction

Two models, depending on the assumption on the number of losses:

- deterministic - n
 - main focus on the claims of **individual policies** (whose number is a priori known)
 - \rightarrow **Individual** Risk Model
 - discussed in previous sections
- random - N
 - main focus on claims of a **whole portfolio** (whose number is a priori unknown)
 - \rightarrow **Collective** Risk Model
 - this is another way of separating frequency and severity

In this section we focus on the Collective Risk Model.

Definition

In the Collective Risk Model, aggregate losses become

$$S = Y_1 + \dots + Y_N = \sum_{i=1}^N Y_i.$$

This is a random sum. We make the following assumptions:

- N is the number of claims
- Y_i is the amount of the i th claim
- the Y_i 's are **iid** with
 - (c)df $G(y)$
 - p(d/m)f $g(y)$
- the Y_i 's and N are mutually independent

3 The Collective Risk Model (Compound distributions, MW 2.1)

- Definition
- Moments of S
- Moment generating function of S
- Distribution of S
- Using R

Moments of S

We have

$$E[S] = E[E[S|N]] = E[NE[Y]] = E[N]E[Y],$$

and

$$\begin{aligned} \text{Var}(S) &= E[\text{Var}(S|N)] + \text{Var}(E[S|N]) \\ &= E[N\text{Var}(Y)] + \text{Var}(E[Y]N) \\ &= E[N]\text{Var}(Y) + E[Y]^2\text{Var}(N) \\ &= E[N](E[Y^2] - E[Y]^2) + E[Y]^2\text{Var}(N) \\ &= E[N]E[Y^2] + E[Y]^2(\text{Var}(N) - E[N]). \end{aligned}$$

3 The Collective Risk Model (Compound distributions, MW 2.1)

- Definition
- Moments of S
- **Moment generating function of S**
- Distribution of S
- Using R

Moment generating function of S

It is possible to get $M_S(t)$ as a function of $M_Y(t)$ and $M_N(t)$:

$$\begin{aligned}M_S(t) &= E \left[e^{tS} \right] = E \left[E \left[e^{t(Y_1 + Y_2 + \dots + Y_N)} \mid N \right] \right] \\&= E \left[M_Y(t)^N \right] = E \left[e^{N \ln M_Y(t)} \right] \\&= M_N(\ln M_Y(t))\end{aligned}$$

Example (Bowers et al. (1997), 12.2.1)

Assume that N is geometric with probability of success p :

$$\Pr[N = n] = pq^n, \quad n = 0, 1, \dots,$$

where $0 < q < 1$ and $p = 1 - q$. We have then

$$M_N(t) = E[e^{tN}] = \sum_{n=0}^{\infty} pq^n e^{tn} = \frac{p}{1 - qe^t},$$

and thus

$$M_S(t) = M_N(\ln M_Y(t)) = \frac{p}{1 - qe^{\ln M_Y(t)}} = \frac{p}{1 - qM_Y(t)}.$$

3 The Collective Risk Model (Compound distributions, MW 2.1)

- Definition
- Moments of S
- Moment generating function of S
- **Distribution of S**
- Using R

Distribution of S

It is possible to get a fairly **general expression for the df of S** by conditioning on the number of claims:

$$F_S(x) = \sum_{n=0}^{\infty} \Pr[S \leq x | N = n] \Pr[N = n] = \sum_{n=0}^{\infty} G^{*n}(x) \Pr[N = n], \quad (1)$$

where $G^{*n}(y)$ is the n -th convolution of G .

Note that

- N will always be discrete, so this works for any type of rv Y . (continuous, discrete or mixed)
- However, the type of S will depend on the type of Y .

Distribution of S if X is continuous

If X is **continuous**, S will generally be mixed:

- with a mass at 0 because of $\Pr[N = 0]$ (if positive)
- continuous elsewhere, but with a density integrating to $1 - \Pr[N = 0]$

Example, continued (Bowers et al. (1997), 12.2.3)

Assume now that

$$G(y) = 1 - e^{-y} \quad \text{and hence} \quad M_Y(t) = \frac{1}{1-t} \quad \text{for } t < 1.$$

Now, we have that (remember $\Pr[N = 0] = p$)

$$M_S(t) = \frac{p}{1 - qM_Y(t)}.$$

It follows that

$$M_S(t) = \frac{p}{1 - q\frac{1}{1-t}} = p + q\frac{p}{p-t} = pE[e^{t \cdot 0}] + (1-p)E[e^{tZ}],$$

where Z is an exponential rv with parameter p . Therefore,

$$f_S(s) = \begin{cases} p = \Pr[N = 0] \text{ (probability mass)} & s = 0; \\ (1-p)(pe^{-ps}) \text{ (probability density)} & s > 0. \end{cases}$$

Distribution of S if Y is mixed

If Y is **mixed**, S will generally be mixed:

- with a mass at 0 because of $\Pr[N = 0]$ and $\Pr[Y = 0]$ (if positive)
- mixed (if Y is not continuous for $x > 0$) or continuous elsewhere
- with a density integrating to something $\leq 1 - \Pr[N = 0]$

Distribution of S if Y is discrete

For discrete Y 's we can get a similar expression to for the pmf of S :

$$f_S(s) = \sum_{n=0}^{\infty} \Pr[S = s | N = n] \Pr[N = n] = \sum_{n=0}^{\infty} g^{*n}(s) \Pr[N = n], \quad (2)$$

where $g^{*0}(0) = 1$ (and thus 0 anywhere else).

- This can be implemented in a table and/or in a program.
- However, if the range of N goes really to infinity, calculating $f_S(s)$ may require an infinity of convolutions of Y .
- This formula is more efficient if the number of possible outcomes for N is small.
- ✂ The pmf $g^{*n}(s)$ can be calculated using de Pril's algorithm. (see Module 4)

Example with tabular approach

From Bowers et al. (1997), 12.2.2:

(1) x	(2) $p^{*0}(x)$	(3) $p^{*1}(x) = p(x)$	(4) $p^{*2}(x)$	(5) $p^{*3}(x)$	(6) $f_S(x)$	(7) $F_S(x)$
0	1.0	—	—	—	0.1000	0.1000
1	—	0.5	—	—	0.1500	0.2500
2	—	0.4	0.25	—	0.2200	0.4700
3	—	0.1	0.40	0.125	0.2150	0.6850
4	—	—	0.26	0.300	0.1640	0.8490
5	—	—	0.08	0.315	0.0950	0.9440
6	—	—	0.01	0.184	0.0408	0.9848
7	—	—	—	0.063	0.0126	0.9974
8	—	—	—	0.012	0.0024	0.9998
9	—	—	—	0.001	0.0002	1.0000
n	0	1	2	3	—	—
$\Pr(N = n)$	0.1	0.3	0.4	0.2	—	—

- The convolutions are in done the usual way.
- The number of columns depends on the range of N .
- The $f_S(x)$ are the sumproduct of the row x and row $\Pr[N = n]$:

3 The Collective Risk Model (Compound distributions, MW 2.1)

- Definition
- Moments of S
- Moment generating function of S
- Distribution of S
- Using R

Using R

We will make extensive use of the function `aggregateDist` from the package `actuar` (Dutang, Goulet, and Pigeon 2008):

- This function allows for several different aggregate distribution approaches, which will be introduced here (and in Module 4 as the associated theory is presented).
- Here, we show how the function can be used to implement formulas (1) and (2) (using the function `convolve` in the background). This corresponds to the `method="convolution"` approach.

```
actuar::aggregateDist(method="convolution"):
```

- A *discrete* distribution for Y is required. Note that discretisation methods are discussed in Module 4. This is input as a vector of claim amount probability *masses* after the argument `model.sev=`. The first element must be $\Pr[Y = 0]$.
- There is no restriction on the shape of the frequency distribution, but it must have a finite range. This is input as a vector of claim number probability masses after the argument `model.freq=`. The first element must be $\Pr[N = 0]$.
- The outcome of the function is (1). Additional outputs:
 - `plot`: to get a pretty plot of the df
 - `summary`: to get summary statistics
 - `mean`: to get the mean
 - `diff`: to get the pmf
- Additional options are:
 - `x.scale`: currency units per unit of sev in the severity model (this allows calculations on multiples of \$1)

```
# Bowers 12.2.2
```

```
fy <- c(0, 0.5, 0.4, 0.1)
```

```
fn <- c(0.1, 0.3, 0.4, 0.2)
```

```
Fs <- aggregateDist("convolution", model.freq = fn, model.sev = fy)
```

```
mean(Fs)
```

```
## [1] 2.72
```

```
pmf <- c(Fs(0), diff(Fs(0:9)))
```

```
cbind(s = c(0:9), fs = pmf, Fs = Fs(0:9))
```

```
##      s      fs      Fs
```

```
## [1,] 0 0.1000 0.1000
```

```
## [2,] 1 0.1500 0.2500
```

```
## [3,] 2 0.2200 0.4700
```

```
## [4,] 3 0.2150 0.6850
```

```
## [5,] 4 0.1640 0.8490
```

```
## [6,] 5 0.0950 0.9440
```

```
## [7,] 6 0.0408 0.9848
```

```
## [8,] 7 0.0126 0.9974
```

```
## [9,] 8 0.0024 0.9998
```

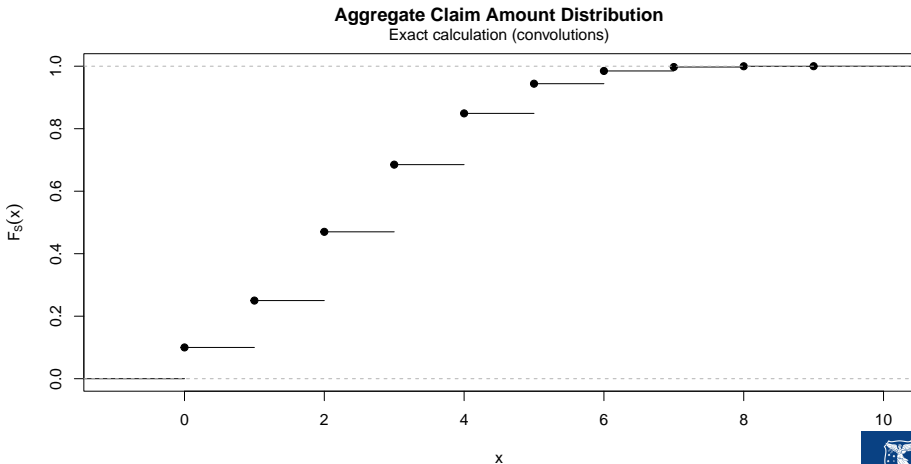
```
## [10,] 9 0.0002 1.0000
```

```
summary(Fs)
```

```
## Aggregate Claim Amount Empirical CDF:
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.00   2.00   3.00   2.72   4.00   9.00
```

```
plot(Fs)
```



- 1 Introduction: Models for aggregate losses
- 2 The Individual Risk Model
- 3 The Collective Risk Model (Compound distributions, MW 2.1)
- 4 Explicit claims count distributions (MW 2.2)**
- 5 Additional properties and applications of Poisson frequencies
- 6 ✕ Parameter estimation (MW 2.3)
- 7 The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions

4 Explicit claims count distributions (MW 2.2)

- Introduction
- Binomial distribution
- Poisson distribution
- Mixed Poisson distribution
- Negative-binomial distribution

Exposure

- It makes no sense to talk about frequency in an insurance portfolio without considering **exposure**. Chapter 4 of Werner and Modlin (2010) defines *exposure* as “the basic unit that measures a policy’s exposure to loss”.
- One primary criterion for choosing an exposure base is that it “should be directly proportional to expected loss”. Here we are focussing on frequency, so exposure should be something directly proportional to the expected frequency.
- Wuthrich (2022) calls exposure “volume”, denoted v , and defines the **claims frequency** as

$$\frac{N}{v}.$$

Basic models for claims frequency

- In our case, we will assume that it directly affects the likelihood of a claim to occur - the frequency - such that N/v is normalised
- MW defines

$$p_k = \Pr[N = k], \quad \text{for } k \in \mathcal{A} \subset \mathbb{N}_0,$$

where \mathcal{A} is the set of possible frequency outcomes.

- There are three main assumptions for p_k :
 - binomial (with variance less than mean)
 - Poisson (with variance equal to the mean)
 - negative-binomial (a Poisson with random mean, so that variance is more than the mean)
- A summary table of those distributions is also given in Bowers et al. (1997), see Table 12.3.1 on page 376.
- These all belong to a class of distributions called (a, b)

4 Explicit claims count distributions (MW 2.2)

- Introduction
- **Binomial distribution**
- Poisson distribution
- Mixed Poisson distribution
- Negative-binomial distribution

Binomial distribution

- fixed volume $v \in \mathbb{N}$
- fixed default probability $p \in (0, 1)$ (expected claims frequency)
- pmf of $N \sim \text{Binom}(v, p)$ is

$$p_k = \Pr[N = k] = \binom{v}{k} p^k (1 - p)^{v-k}, \quad \text{for all } k \in \{0, \dots, v\} = \mathcal{A}.$$

- same as a sum of Bernoulli (which is the case $v = 1$)
- makes sense for homogenous portfolio with unique possible events, such as credit defaults, or deaths in a life insurance model
- In R: `dbinom`, `pbinom`, `qbinom`, `rbinom`, where `size` is v , and where `prob` is p
- Note that $\binom{v}{k}$ can be computed with the R function `choose`.

Compound binomial model

The total claim amount S has a **compound binomial distribution**

$$S \sim \text{CompBinom}(v, p, G)$$

if S has a compound distribution with $N \sim \text{Binom}(v, p)$ for given $v \in \mathbb{N}$ and $p \in (0, 1)$ and individual claim size distribution G .

Corollary 2.7: Assume S_1, \dots, S_n are independent with $S_j \sim \text{CompBinom}(v_j, p, G)$ for all $j = 1, \dots, n$. The aggregated claim has a compound binomial distribution with

$$S = \sum_{j=1}^n S_j \sim \text{CompBinom} \left(\sum_{j=1}^n v_j, p, G \right).$$

Exercise NLI3 considers the decomposition of S into small and large claims. It shows that S_{lc} —the sum of those claims exceeding a certain threshold M only (see notation in Example 2.16 later in those slides)—is compound binomial again.

4 Explicit claims count distributions (MW 2.2)

- Introduction
- Binomial distribution
- **Poisson distribution**
- Mixed Poisson distribution
- Negative-binomial distribution

Poisson distribution

- fixed volume $v > 0$
- expected claims frequency $\lambda > 0$
- pmf of $N \sim \text{Poi}(\lambda v)$ is

$$p_k = \Pr[N = k] = e^{-\lambda v} \frac{(\lambda v)^k}{k!} \quad \text{for all } k \in \mathcal{A} = \mathbb{N}_0.$$

- Lemma 2.9: increase volume while keeping $E[N]$ fixed in a binomial model leads to a Poisson distribution (more so for small p compared to v).
- In R: `dpois`, `ppois`, `qpois`, `rpois`, where `lambda` is λv

Compound Poisson model

The total claim amount S has a **compound Poisson distribution**

$$S \sim \text{CompPoi}(\lambda\nu, G)$$

if S has a compound distribution with $N \sim \text{Poi}(\lambda\nu)$ for given $\lambda, \nu > 0$ and individual claim size distribution G .

- The compound Poisson distribution has nice properties such as:
 - The aggregation property \uparrow
 - The disjoint decomposition property \downarrow
- These are reviewed in the next section, along with related new techniques for computing the distribution of S .

4 Explicit claims count distributions (MW 2.2)

- Introduction
- Binomial distribution
- Poisson distribution
- **Mixed Poisson distribution**
- Negative-binomial distribution

Inhomogeneous portfolio

- So far we have seen distributions with variance less (binomial) or exactly equal (Poisson) to the mean.
- In reality, actuarial data is often **overdispersed**, that is, variance is larger than mean.
- This could be due to frequency or severity, but it makes sense that some of this extra variability would come from frequency.
- If we believe in a Poisson frequency *for known frequency parameter*, then additional uncertainty such as heterogeneity of risks in a portfolio, uncertain conditions (weather, for instance) could be modelled with a **random** Poisson parameter, and could explain the extra variability.
- This is the idea of a mixed Poisson.

The mixed Poisson distribution

- Assume **random** $\Lambda \sim H$ with $H(0) = 0$, $E[\Lambda] = \lambda$, and $\text{Var}(\Lambda) > 0$.
- Conditionally, given Λ , $N \sim \text{Poi}(\Lambda v)$ for fixed volume $v > 0$.

We have then

$$\Pr[N = n] = \int_0^\infty \Pr[N = n | \Lambda = \lambda] dH(\lambda) = \int_0^\infty \frac{e^{-\lambda v} (\lambda v)^n}{n!} dH(\lambda);$$

$$E[N] = E[E[N | \Lambda]] = E[\Lambda]v = \lambda v;$$

$$\text{Var}(N) = E[\text{Var}(N | \Lambda)] + \text{Var}(E[N | \Lambda]) = \lambda v + v^2 \text{Var}(\Lambda) > \lambda v;$$

$$M_N(t) = E[e^{tN}] = E[E[e^{tN} | \Lambda]] = E[e^{\Lambda v(e^t - 1)}] = M_\Lambda(v[e^t - 1]).$$

Example

If $\Lambda \sim \text{inverse Gaussian}(\alpha, \beta)$ (Example 12.3.2):

- N is Poisson Inverse Gaussian.
- This distribution is the `pig` distribution in `actuar`, so that you can use `dpig`, `ppig`, etc...); see Section 5 of the vignette “distribution” of `actuar`.
- $\rightarrow S$ will be compound inverse Gaussian.

Another example, which is very famous, is $\Lambda \sim \Gamma$, which leads to the negative-binomial distribution.

4 Explicit claims count distributions (MW 2.2)

- Introduction
- Binomial distribution
- Poisson distribution
- Mixed Poisson distribution
- Negative-binomial distribution

Negative-binomial distribution

Assume λ is the mean, and will be “spread” according to a gamma distribution:

- Define $\Lambda = \lambda\Theta$.
- Now, $\Theta \sim \Gamma(\gamma, \gamma)$ such that

$$E[\Theta] = 1 \quad \text{and} \quad \text{Var}(\Theta) = \frac{1}{\gamma}$$

and

$$E[\Lambda] = \lambda \quad \text{and} \quad \text{Var}(\Lambda) = \frac{\lambda^2}{\gamma}.$$

- If conditionally, given Θ , $N \sim \text{Poi}(\Theta\lambda\nu)$, then

$$N \sim \text{NegBin}(\lambda\nu, \gamma)$$

with volume $\nu > 0$, expected claims frequency $\lambda > 0$,
and dispersion parameter $\gamma > 0$.

Proof:

$$\begin{aligned}
 M_N(t) &= E[e^{tN}] = E[E[e^{tN}|\Theta]] = E[e^{\Theta\lambda v(e^t-1)}] \\
 &= \left(\frac{\gamma}{\gamma-\lambda v(e^t-1)}\right)^\gamma = \left(\frac{\gamma}{\gamma+\lambda v-\lambda ve^t}\right)^\gamma = \left(\frac{\frac{\gamma}{\lambda v+\gamma}}{1-\frac{\lambda v}{\lambda v+\gamma}e^t}\right)^\gamma,
 \end{aligned}$$

which can be recognised as a negative-binomial with probability of “failure”

$$p = \frac{\lambda v}{\lambda v + \gamma}$$

(if we count failures until the γ -th success) so that

$$p_k = \Pr[N = k] = \binom{k + \gamma - 1}{k} p^\gamma (1 - p)^k$$

In R, use `dnbinom`, `pnbinom`, `qnbinom`, `rnbinom`, where `size` is γ and `prob` is probability of success $1 - p$ (note volume is hidden in p and will affect the scale of the distribution).

Interpretation

- Θ reflects the uncertainty about the ‘true’ parameter of the Poisson distribution.

Alternatively, it describes the distributions of “ λ ’s” in the population.

- In the end we have

$$\begin{aligned} E[N] &= \lambda v, \\ \text{Var}(N) &= \lambda v \left(1 + \frac{\lambda v}{\gamma}\right) > \lambda v, \\ \text{Vco}\left(\frac{N}{v}\right) &= \sqrt{(\lambda v)^{-1} + \gamma^{-1}}. \end{aligned}$$

- This additional uncertainty is not diversifiable (remains even for large v):

$$\text{Vco}\left(\frac{N}{v}\right) = \sqrt{(\lambda v)^{-1} + \gamma^{-1}} \rightarrow \gamma^{-1/2} > 0 \quad \text{for } v \rightarrow \infty.$$

Compound negative-binomial model

The total claim amount S has a **compound negative-binomial distribution**

$$S \sim \text{CompNB}(\lambda\nu, \gamma, G)$$

if S has a compound distribution with $N \sim \text{NegBin}(\lambda\nu, \gamma)$ for given $\lambda, \nu, \gamma > 0$ and individual claim size distribution G .

- 1 Introduction: Models for aggregate losses
- 2 The Individual Risk Model
- 3 The Collective Risk Model (Compound distributions, MW 2.1)
- 4 Explicit claims count distributions (MW 2.2)
- 5 Additional properties and applications of Poisson frequencies**
- 6 ✕ Parameter estimation (MW 2.3)
- 7 The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions

5 Additional properties and applications of Poisson frequencies

- Theorem 2.12: Aggregation property \uparrow
- Theorem 2.14: Disjoint decomposition property \downarrow
- Thinning of the Poisson process
- \boxtimes Sparse vector algorithm
- Example 2.16: Large claim separation

Theorem 2.12: Aggregation property ↑

Assume S_1, \dots, S_n are independent with $S_j \sim \text{CompPoi}(\lambda_j v_j, G_j)$ for all $j = 1, \dots, n$. Aggregated claims have a compound Poisson distribution

$$S = \sum_{j=1}^n S_j \sim \text{CompPoi}(\lambda v, G), \text{ with}$$

$$v = \sum_{j=1}^n v_j, \quad \lambda = \sum_{j=1}^n \frac{v_j}{v} \lambda_j, \quad G = \sum_{j=1}^n \frac{\lambda_j v_j}{\lambda v} G_j.$$

So what?

- Independent n portfolios of losses can be easily aggregated.
- Alternatively (or in addition), total claims paid over n years are compound Poisson, even if the severity and frequency of losses vary across years.
- “Bottom-up” modelling
- In Bowers et al. (1997), this is Theorem 12.4.1.

Example 12.4.1 of Bowers et al. (1997)

Suppose that N_1, N_2, \dots, N_m are independent random variables. Further, suppose that N_i follows $\text{Poisson}(\lambda_i)$. Let y_1, y_2, \dots, y_m be deterministic numbers. What is the distribution of

$$y_1 N_1 + \dots + y_m N_m?$$

5 Additional properties and applications of Poisson frequencies

- Theorem 2.12: Aggregation property ↑
- Theorem 2.14: Disjoint decomposition property ↓
- Thinning of the Poisson process
- ✂ Sparse vector algorithm
- Example 2.16: Large claim separation

Preliminary 1: Add LoBs in the CompPoi formulation

Let us introduce Lines of Business (“LoB”) in the notation:

- Let the set $\{1, \dots, m\}$ be a partition of the portfolio, or different lines of business (“LoB” thereafter). For instance, we could have $j \in \{1, 2, 3\}$ for car ($j = 1$), building ($j = 2$) and liability ($j = 3$) LoBs.
- Let $(p_j^+)_{j=1, \dots, m}$ be a discrete probability distribution on the finite set of sub-portfolios/LoBs $\{1, \dots, m\}$ (thereafter just “LoB”).
- We assume

$$p_j^+ > 0 \text{ for all } j,$$

that is, the probability of having claims in any of the m LoBs is strictly positive.

- We further assume that G_j is the claim size distribution of LoB j , with $G_j(0) = 0$.

- Finally, we define the mixture distribution by

$$G(y) = \sum_{j=1}^m p_j^+ G_j(y) \quad \text{for } y \in \mathbb{R}.$$

This is the distribution of a claim, if we don't know which LoB it comes from.

- Note that this matches the formulation in the aggregation property Theorem 2.12 with

$$p_j^+ = \frac{\lambda_j v_j}{\lambda_V}.$$

- Now, define a discrete random variable I which indicates which sub-portfolio/LoB a randomly selected claim Y belongs to:

$$\Pr[I = j] = p_j^+ \quad \text{for all } j \in \{1, \dots, m\}.$$

We are now ready to define the following extended compound Poisson model:

- The total claims $S = \sum_{i=1}^N Y_i$ has a compound Poisson distribution as defined earlier.
- In addition, we assume that

$$(Y_i, I_i)_{i \geq 1}$$

are

- mutually i.i.d. and independent of N ,
- with Y_i having marginal distribution function G with $G(0) = 0$, and
- I_i having marginal distribution function given by $\Pr[I = j] = p_j^+$ for all $j \in \{1, \dots, n\}$.

Preliminary 2: Partition

- The random vector (Y_1, I_1) takes values in $\mathbb{R}_+ \times \{1, \dots, m\}$.
- On this set we choose a finite sequence of sets

$$A_1, \dots, A_n$$

such that

$$\begin{aligned} A_k \cap A_l &= \emptyset \text{ for all } k \neq l \quad (\text{no overlap}); \\ \bigcup_{k=1}^n A_k &= \mathbb{R}_+ \times \{1, \dots, m\} \quad (\text{all-inclusive}). \end{aligned}$$

- Such a sequence is called a “measurable disjoint decomposition” or “**partition**” of $\mathbb{R}_+ \times \{1, \dots, m\}$.
- This partition is called “admissible” for (Y_1, I_1) if for all $k = 1, \dots, n$

$$p^{(k)} = \Pr[(Y_1, I_1) \in A_k] > 0.$$

Note $\sum_{k=1}^n p^{(k)} = 1$ due to the properties of the partition above (no overlap and all-inclusive)

We have two levels of partition:

- Into LoBs:
 - Claims are classified according to a sub-portfolio or LoB
 - For instance: domestic motor and commercial motor
 - The probability of a claim being in LoB j is p_j^+
 - The indicator for the claim to be in LoB j is I_j
(with probability p_j^+ of being 1)
- Into a second level:
 - Claims are classified according to another set of criteria
 - For instance: geographical areas NSW and VIC
 - The probability of a claim being in geographical area k is $p(k)$

Theorem 2.14: Disjoint decomposition ↓

Assume S is “doubly partitioned” as described above:

- S fulfills the extended compound Poisson model assumptions above (Preliminary 1).
- We chose an admissible partition A_1, \dots, A_n for (Y_1, I_1) (Preliminary 2).

Then the random variable (sum of claims for partition k):

$$S_k = \sum_{i=1}^N Y_i 1_{\{(Y_i, I_i) \in A_k\}} \sim \text{CompPoi}(\lambda_k v_k, G_k),$$

for $k = 1, \dots, n$, with

$$\lambda_k v_k = \lambda v p^{(k)} > 0, \quad G_k(y) = \Pr[Y_1 \leq y | (Y_1, I_1) \in A_k].$$

Furthermore, the S_k 's are independent (over k).

5 Additional properties and applications of Poisson frequencies

- Theorem 2.12: Aggregation property \uparrow
- Theorem 2.14: Disjoint decomposition property \downarrow
- Thinning of the Poisson process
- \boxtimes Sparse vector algorithm
- Example 2.16: Large claim separation

Thinning of the Poisson process

- Assume that $m = 1$ (only one LoB)
- The disjoint decomposition theorem implies that

$$Y_i = Y_i 1_{\{Y_i \in A_1\}} + \dots + Y_i 1_{\{Y_i \in A_n\}}.$$

- For each partition A_k (defined on the claims) a natural choice is
 - $v_k = v$
 - $\lambda_k = \lambda p^{(k)}$
- This means that the volume remains constant in each partition, but the expected claims frequencies λ_k change proportionally to the probabilities of falling in partition A_k , $k = 1, \dots, n$.
- This is called **thinning** of the Poisson process.

5 Additional properties and applications of Poisson frequencies

- Theorem 2.12: Aggregation property \uparrow
- Theorem 2.14: Disjoint decomposition property \downarrow
- Thinning of the Poisson process
- ✂ Sparse vector algorithm
- Example 2.16: Large claim separation

✂ Sparse vector algorithm

If $S \sim \text{compound Poisson}(\lambda, g(y_i) = \pi_i)$, $i = 1, \dots, m$ then

$$S = y_1 N_1 + \dots + y_m N_m,$$

where the N_i 's

- represent the number of claims of amount y_i ;
- are mutually independent;
- are $\text{Poi}(\lambda_i = \lambda \pi_i)$.

Proof: see tutorial exercise 1os18. Note also that this is a special case of Theorem 2.14, and is Theorem 12.4.2 of Bowers et al. (1997).

So what?

- **Sparse vector algorithm:** allows to develop an alternative method for tabulating the distribution of S that is more efficient as m is small.
- S can be used to approximate the Individual Risk Model if $X = Ib$ (see Module 3).

✖ The sparse vector algorithm

(Bowers et al. 1997, Example 12.4.2) Suppose S has a compound Poisson distribution with $\lambda = 0.8$ and individual claim amount distribution

y_i	$\Pr[Y = y_i]$
1	0.250
2	0.375
3	0.375

Compute $f_S(s) = \Pr[S = s]$ for $s = 0, 1, \dots, 6$.

This can be done in two ways:

- Basic method (seen earlier in the lecture): requires to calculate up to the 6th convolution of Y .
- Sparse vector algorithm: requires no convolution of Y .

Solution - Basic Method

x	$g^{*0}(x)$	$g(x)$	$g^{*2}(x)$	$g^{*3}(x)$	$g^{*4}(x)$	$g^{*5}(x)$	$g^{*6}(x)$	$f_S(x)$
0	1	-	-	-	-	-	-	0.4493
1	-	0.250	-	-	-	-	-	0.0899
2	-	0.375	0.0625	-	-	-	-	0.1438
3	-	0.375	0.1875	0.0156	-	-	-	0.1624
4	-	-	0.3281	0.0703	0.0039	-	-	0.0499
5	-	-	0.2813	0.1758	0.0234	0.0010	-	0.0474
6	-	-	0.1406	0.2637	0.0762	0.0073	0.0002	0.0309
n	0	1	2	3	4	5	6	
$\Pr[N = n]$	0.4493	0.3595	0.1438	0.0383	0.0077	0.0012	0.0002	
$= \frac{e^{-0.8} (0.8)^n}{n!}$								

- The convolutions are done in the usual way.
- The $f_S(x)$ are the sumproduct of the row x and row $\Pr[N = n]$.
- The number of convolutions (and thus of columns) will increase by 1

Solution - Sparse vector algorithm

Thanks to Theorem 2.12, we can write $S = N_1 + 2N_2 + 3N_3$

x	$\Pr[N_1 = x]$	$\Pr[2N_2 = x]$	$\Pr[3N_3 = x]$	$\Pr[N_1 + 2N_2 + 3N_3 = x]$	$f_S(x)$
0	0.818731	0.740818	0.740818	0.606531	0.449329
1	0.163746	0	0	0.121306	0.089866
2	0.016375	0.222245	0	0.194090	0.143785
3	0.001092	0	0.222245	0.037201	0.162358
4	0.000055	0.033337	0	0.030974	0.049906
5	0.000002	0	0	0.005703	0.047360
6	0.000000	0.003334	0.033337	0.003288	0.030923
x_i	1	2	3		
λ_i	0.2	0.3	0.3		
$\lambda \pi_i$					
$\Pr[N_i = x/i]$	$e^{-0.2} \frac{(0.2)^x}{x!}$	$e^{-0.3} \frac{(0.3)^{x/2}}{(x/2)!}$	$e^{-0.3} \frac{(0.3)^{x/3}}{(x/3)!}$		

5 Additional properties and applications of Poisson frequencies

- Theorem 2.12: Aggregation property \uparrow
- Theorem 2.14: Disjoint decomposition property \downarrow
- Thinning of the Poisson process
- \boxplus Sparse vector algorithm
- Example 2.16: Large claim separation

Example 2.16: Large claim separation

- This is a very important (and convenient) application of the Disjoint decomposition property (Theorem 2.14).
- Attritional and catastrophic claims often have very different distributions (different G 's); see also <https://www.actuaries.digital/2022/01/10/catastrophe-vs-standard-loss-modelling/>
- The idea here is to divide the claims into different layers with different distributions:
 - Small claims are modelled using a parametric distribution for which it is easy to obtain the distribution of the compound distribution, potentially even approximated with a normal distribution thanks to volume and light right tail;
 - Large claims are typically modelled with a Pareto distribution with threshold M and tail parameter $\alpha > 1$ (see Module 6 for a justification of this, and for the choice of an appropriate M). The could also be “modelled” (see article above)

Assuming two layers:

- We choose a large claims threshold $M > 0$ such that

$$0 < G(M) < 1,$$

that is, there is probability mass on either size of M .

- We define the partition

$$A = A_{\text{sc}} = \{Y_1 \leq M\} \quad \text{and} \quad A^c = A_{\text{lc}} = \{Y_1 > M\}.$$

- Assume that

$$S \sim \text{CompPoi}(\lambda v, G).$$

- We now define the small and large claims layers as

$$\begin{aligned} S_{\text{sc}} &= \sum_{i=1}^N Y_i 1_{\{Y_i \leq M\}}, \text{ and} \\ S_{\text{lc}} &= \sum_{i=1}^N Y_i 1_{\{Y_i > M\}}, \end{aligned}$$

respectively.

- Theorem 2.14 implies that S_{sc} and S_{lc} are *independent* and compound Poisson distributed with

$$\begin{aligned} S_{sc} &\sim \text{CompPoi}(\lambda_{sc}v = \lambda G(M)v, \\ &\quad G_{sc}(y) = \Pr[Y_1 \leq y | Y_1 \leq M]), \text{ and} \\ S_{lc} &\sim \text{CompPoi}(\lambda_{lc}v = \lambda(1 - G(M))v, \\ &\quad G_{lc}(y) = \Pr[Y_1 \leq y | Y_1 > M]), \end{aligned}$$

respectively.

- The distribution of

$$S = S_{sc} + S_{lc}$$

can then be obtained by a simple convolution of distributions of S_{sc} and S_{lc} (thanks to independence); see Module 4 for examples (✂).

- 1 Introduction: Models for aggregate losses
- 2 The Individual Risk Model
- 3 The Collective Risk Model (Compound distributions, MW 2.1)
- 4 Explicit claims count distributions (MW 2.2)
- 5 Additional properties and applications of Poisson frequencies
- 6 ✠ Parameter estimation (MW 2.3)**
- 7 The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions

6 ✦ Parameter estimation (MW 2.3)

- ✦ Introduction
- ✦ What to do with volumes? Lemma 2.26
- ✦ Method of moments
- ✦ Maximum likelihood estimators

✦ Estimation methods

You should be familiar with the main estimation methods:

- Method of moments
- Maximum likelihood estimation

Here the problem is slightly complicated because our observations may not be directly comparable due to varying exposures v 's.

Assume that $(N_1, \dots, N_T)'$ is the vector of observations.

6 ✦ Parameter estimation (MW 2.3)

- ✦ Introduction
- ✦ What to do with volumes? Lemma 2.26
- ✦ Method of moments
- ✦ Maximum likelihood estimators

✦ What to do with volumes? Lemma 2.26

- The key idea here is to find the minimum variance method of moments estimator, when the volumes across the observations can vary.
- This is what is different from a straight method of moments estimator, and explains why we need to think it through: how to deal with those volumes?
- Assume there exist strictly positive volumes v_1, \dots, v_T such that the components of $(N_1/v_1, \dots, N_T/v_T)$ are independent with

$$\lambda = E \left[\frac{N_t}{v_t} \right] \text{ and } \tau_t^2 = \text{Var} \left(\frac{N_t}{v_t} \right) \in (0, \infty),$$

for all $t = 1, \dots, T$.

Lemma 2.26 states that the unbiased, linear estimator for λ with minimal variance is given by

$$\hat{\lambda}_T^{\text{MV}} = \left(\sum_{t=1}^T \frac{1}{\tau_t^2} \right)^{-1} \sum_{t=1}^T \frac{N_t/v_t}{\tau_t^2},$$

with variance

$$\text{Var}(\hat{\lambda}_T^{\text{MV}}) = \left(\sum_{t=1}^T \frac{1}{\tau_t^2} \right)^{-1}.$$

Note:

- We haven't made any distributional assumption yet - this estimates $E \left[\frac{N_t}{v_t} \right]$ via method of moments, taking the v_t 's into account in an optimal way (in the sense that it minimises the variance of the estimator).
- The superscript “MV” stands for “minimal variance”.

6 ✦ Parameter estimation (MW 2.3)

- ✦ Introduction
- ✦ What to do with volumes? Lemma 2.26
- ✦ Method of moments
- ✦ Maximum likelihood estimators

✂ Binomial and Poisson cases

Unbiased, minimal variance estimators:

- binomial case for p :

$$\hat{p}_T^{\text{MV}} = \frac{1}{\sum_{s=1}^T v_s} \sum_{t=1}^T N_t = \sum_{t=1}^T \frac{v_t}{\sum_{s=1}^T v_s} \frac{N_t}{v_t} \sim$$

Furthermore, $\sum_{t=1}^T N_t \sim \text{Binom}(\sum_{s=1}^T v_s, p)$, which means we know the distribution of \hat{p}_T^{MV} .

- Poisson case for λ :

$$\hat{\lambda}_T^{\text{MV}} = \frac{1}{\sum_{s=1}^T v_s} \sum_{t=1}^T N_t = \sum_{t=1}^T \frac{v_t}{\sum_{s=1}^T v_s} \frac{N_t}{v_t}$$

Here, $\sum_{t=1}^T N_t \sim \text{Poi}(\lambda \sum_{s=1}^T v_s)$.

✦ Negative binomial case

More complicated, because:

$$E \left[\frac{N_t}{v_t} \right] = \lambda \text{ and } Var \left(\frac{N_t}{v_t} \right) = \lambda/v_t + \lambda^2/\gamma = \tau_t^2,$$

Unbiased (but not guaranteed minimal variance):

$$\hat{\lambda}_T^{\text{NB}} = \frac{1}{\sum_{s=1}^T v_s} \sum_{t=1}^T N_t = \sum_{t=1}^T \frac{v_t}{\sum_{s=1}^T v_s} \frac{N_t}{v_t}$$

✦ We need a sense of the dispersion for estimating the dispersion parameter γ .

Let the weighted sample variance

$$\hat{V}_T^2 = \frac{1}{T-1} \sum_{t=1}^T v_t \left(\frac{N_t}{v_t} - \hat{\lambda}_T^{\text{NB}} \right)^2.$$

Then we have

$$\hat{\gamma}_T^{\text{NB}} = \frac{(\hat{\lambda}_T^{\text{NB}})^2}{\hat{V}_T^2 - \hat{\lambda}_T^{\text{NB}}} \frac{1}{T-1} \left(\sum_{t=1}^T v_t - \frac{\sum_{t=1}^T v_t^2}{\sum_{t=1}^T v_t} \right),$$

ONLY if $\hat{V}_T^2 > \hat{\lambda}_T^{\text{NB}}$. Otherwise use Poisson or binomial.

6 ✦ Parameter estimation (MW 2.3)

- ✦ Introduction
- ✦ What to do with volumes? Lemma 2.26
- ✦ Method of moments
- ✦ Maximum likelihood estimators

✦ Binomial and Poisson cases

Estimators are identical to method of moments estimators. Or conversely, the MLE estimators are actually unbiased.

- binomial case for p :

$$\hat{p}_T^{\text{MLE}} = \frac{1}{\sum_{s=1}^T v_s} \sum_{t=1}^T N_t = \sum_{t=1}^T \frac{v_t}{\sum_{s=1}^T v_s} \frac{N_t}{v_t} = \hat{p}_T^{\text{MV}}$$

- Poisson case for λ :

$$\hat{\lambda}_T^{\text{MLE}} = \frac{1}{\sum_{s=1}^T v_s} \sum_{t=1}^T N_t = \sum_{t=1}^T \frac{v_t}{\sum_{s=1}^T v_s} \frac{N_t}{v_t} = \hat{\lambda}_T^{\text{MV}}$$

✦ Negative binomial case

Assume N_1, \dots, N_T are independent and $\text{NegBin}(\lambda v_t, \gamma)$. The MLE $(\hat{\lambda}_T^{\text{MLE}}, \hat{\gamma}_T^{\text{MLE}})$ are the solution of

$$\frac{\partial}{\partial(\lambda, \gamma)} \sum_{t=1}^T \log \binom{N_t + \gamma - 1}{N_t} + \gamma \log(1 - p_t) + N_t \log p_t = 0,$$

with $p_t = \lambda v_t / (\gamma + \lambda v_t) \in (0, 1)$.

- 1 Introduction: Models for aggregate losses
- 2 The Individual Risk Model
- 3 The Collective Risk Model (Compound distributions, MW 2.1)
- 4 Explicit claims count distributions (MW 2.2)
- 5 Additional properties and applications of Poisson frequencies
- 6 ✕ Parameter estimation (MW 2.3)
- 7 The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions

- 7 The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions
- The (a, b) class of Panjer distributions (4.2.1)
 - ✖ actuar and the $(a, b, 1)$ class
 - ✖ The $(a, b, 1)$ class of distributions

The (a, b) class of Panjer distributions (4.2.1)

A class of distributions has the following property

$$\Pr[N = n] = \left(a + \frac{b}{n}\right) \Pr[N = n - 1], \text{ or } \frac{p_k}{p_{k-1}} = \left(a + \frac{b}{k}\right).$$

This is the (a, b) class of “Panjer distributions”. This means that $\Pr[N = n]$ can be obtained recursively with initial value $\Pr[N = 0]$; see Wuthrich (2022), Definition 4.6.

The exhaustive list of its members (see Wuthrich 2022 Lemma 4.7) is

Distribution	a	b	$\Pr[N = 0]$
Poisson (λ)	0	λ	$e^{-\lambda}$
Neg Bin (r, p)	$1 - p$	$(r - 1)(1 - p)$	p^r
Binomial (m, p)	$-p/(1 - p)$	$(m + 1)p/(1 - p)$	$(1 - p)^m$

Exercise: prove the results in the above table! (Note p is the probability of success for the Negative Binomial)

First three cumulants of the (a, b) family

Distribution	$E[N]$	$Var(N)$	$E[(N - E[N])^3]$
Poisson (λ)	λ	λ	λ
Neg Bin (r, p)	$\frac{rq}{p}$	$\frac{rq}{p^2}$	$\frac{rq(1+q)}{p^3}$
Binomial (m, p)	mp	mpq	$mpq(q-p)$

Exercise:

- check these results using the cgf
- find the first 3 cumulants of S , as well as ς_S for each member of the family

- 7 The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions
- The (a, b) class of Panjer distributions (4.2.1)
 - ✕ actuar and the $(a, b, 1)$ class
 - ✕ The $(a, b, 1)$ class of distributions

✖ actuar and the $(a, b, 1)$ class

- The package `actuar` extends the definition above to allow for **zero-truncated** and **zero-modified** distributions.
- The Poisson, binomial and negative-binomial (and special case geometric) are all well supported in Base R with the `d`, `p`, `q` and `r` functions.
- If one takes the Panjer equation for granted, then we can think of p_0 as the mass that will make the pmf add up to one:

$$\text{given Panjer : } p_0 \text{ is such that } \sum_{k=0}^{\infty} p_k = 1.$$

- We introduce here the $(a, b, 1)$ class which extends the idea above so that we have more freedom on the mass at 0.
- The reference for this section is Section 4 of the vignette “distribution” of `actuar`

- 7 The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions
- The (a, b) class of Panjer distributions (4.2.1)
 - ✖ actuar and the $(a, b, 1)$ class
 - ✖ The $(a, b, 1)$ class of distributions

✖ The $(a, b, 1)$ class of distributions

A discrete random variable is a member of the $(a, b, 1)$ class of distributions if there exist constants a and b such that

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k}, \quad k = 2, 3, \dots$$

Note:

- The recursion starts at $k = 2$ for the $(a, b, 1)$ class.
- The extra freedom allows the probability at zero to be set to any arbitrary number $0 \leq p_0 \leq 1$

✚ Zero-truncated distributions

- Setting $p_0 = 0$ in the $(a, b, 1)$ class defines the subclass of **zero-truncated distributions**
- Members are the zero-truncated Poisson (`actuar::ztpois`), zero-truncated binomial (`actuar::ztbinom`), zero-truncated negative-binomial (`actuar::ztnbinom`), and the zero-truncated geometric (`actuar::ztgeom`).
- Let p_k^T denote the probability mass at k for a zero-truncated distribution (“ T ” for truncated). We have

$$p_k^T = \begin{cases} 0, & k = 0; \\ \frac{p_k}{1-p_0}, & k = 1, 2, \dots \end{cases},$$

where p_k is the probability mass of the corresponding member of the $(a, b, 0)$ — that is, (a, b) — class.

- `actuar` provides the `d`, `p`, `q`, and `r` functions of the zero-truncated distributions mentioned above.

✂ Zero-modified distributions

- Setting $p_0 \equiv p_0^M$ ($0 < p_0^M < 1$) in the $(a, b, 1)$ class defines the subclass of **zero-modified distributions** (“M” for “modified”)
- These distributions are discrete mixtures between a degenerate distribution at zero, and the corresponding distribution from the $(a, b, 0)$ class.
- Let p_k^M denote the probability mass at k for a zero-modified distribution. We have then

$$p_k^M = \left(1 - \frac{1 - p_0^M}{1 - p_0}\right) 1_{\{k=0\}} + \frac{1 - p_0^M}{1 - p_0} p_k.$$

Alternatively,

$$p_k^M = \begin{cases} p_0^M, & k = 0; \\ \frac{1 - p_0^M}{1 - p_0} p_k, & k = 1, 2, \dots \end{cases},$$

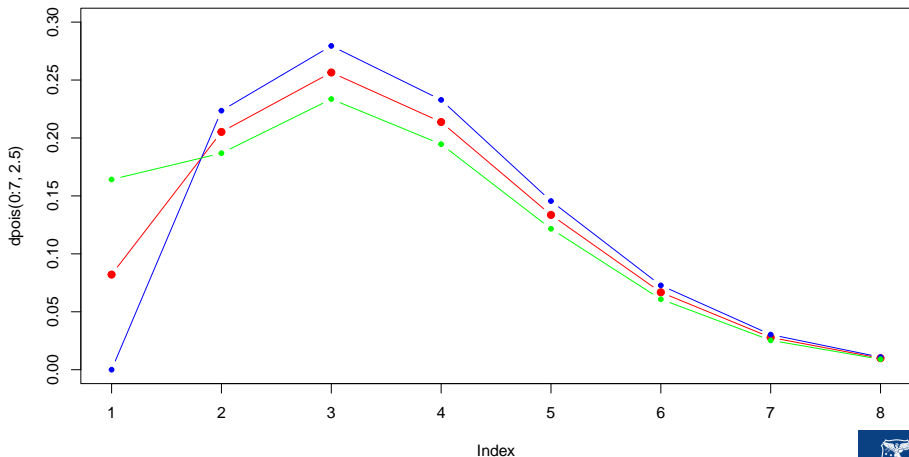
where p_k is the probability mass of the corresponding member of the $(a, b, 0)$ class.

- Quite obviously, zero-truncated distributions are zero-modified distributions with $p_0^M = 0$, and

$$p_k^M = p_0^M 1_{\{k=0\}} + (1 - p_0^M) p_k^T.$$

- Members are the zero-modified Poisson (`actuar::zmpois`), zero-modified binomial (`actuar::zmbinom`), zero-modified negative-binomial (`actuar::zmnbinom`), and the zero-modified geometric (`actuar::zmgeom`). `actuar` provides the `d`, `p`, `q`, and `r` functions of the zero-truncated distributions mentioned above.

```
plot(dpois(0:7, 2.5), pch = 20, col = "red", ylim = c(0, 0.3),  
     cex = 1.5, type = "b")  
points(dztpois(0:7, 2.5), pch = 20, col = "blue", type = "b")  
points(dzmpois(0:7, 2.5, 2 * dpois(0, 2.5)), pch = 20, col = "green",  
       type = "b")
```



References I

- Bowers, Newton L. Jr, Hans U. Gerber, James C. Hickman, Donald A. Jones, and Cecil J. Nesbitt. 1997. *Actuarial Mathematics*. Second. Schaumburg, Illinois: The Society of Actuaries.
- Dutang, Christophe, Vincent Goulet, and Mathieu Pigeon. 2008. "Actuar: An r Package for Actuarial Science." *Journal of Statistical Software* 25 (7).
- Werner, Geoff, and Claudine Modlin. 2010. *Basic Ratemaking*. Casualty Actuarial Society.
- Wuthrich, Mario V. 2022. "Non-Life Insurance: Mathematics & Statistics." Lecture notes. RiskLab, ETH Zurich; Swiss Finance Institute.