ISEG - Lisbon School of Economics and Management

Risk Theory
Master in Actuarial Science
Instructor: Carlos Oliveira

March 10, 2021

1 Risk and Reinsurance Companies

Dealing with risk is the daily-life of many decision-makers in the financial industry. All types of financial decisions are subject to risk, either speculative risk or pure risk. There is a speculative risk whenever a decision may provide losses or profit. This type of risk is generally not insurable. On the other hand, there is a pure risk when an activity can provide losses (or not), but there is no gain in this activity. Usually, these are the insurable risks.

Historically, insurance companies are recognized for their proficiency in measuring the risk of insurable activities. To measure the risk, insurance companies use probabilistic and statistical models. The development of such models started (actuarial mathematics) began in the seventeen century with the first mortality table of Sir Edmund Halley.

This course aims to study and develop mathematical models that correctly describe the technical aspects of the insurance business.

2 The number of claims

In this chapter, we intend to present the necessary mathematical instruments used to model the number of claims. Here, we will not distinguish between claims and losses.

The claim number processes are by its nature counting processes. When the time is fixed we have a counting distribution. If N represents the number of claims that happen in a fixed period of time, for a given risk, then it can be characterized by the following functions:

- Probability function: $p_k = \Pr\{N = k\},\ k = 0, 1, 2, ...$
- Probability generating function: $P_N(z) = E[z^N]$

- Moment generating function: $M_N(r) = E[e^{rN}]$
- Cumulant generating function: $g_N(s) = \ln(M_N(r)) = \sum_{k=1}^{\infty} \kappa_k \frac{s^k}{k!}$.

2.1 The $(a, b, 0)$ class of distributions

Before we start presenting the $(a, b, 0)$ class of distributions, we introduce the Poisson, the Negative Binomial and Binomial distributions.

2.1.1 The Poisson distribution

N is a Poisson random variable if its probability functions is given by

$$p_k = \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!}, \ k = 0, 1, 2, ...$$

Additionally, the probability generating function, the moment generating function and the cumulant generating function are respectively

$$P_N(z) = e^{\lambda(z-1)}, \quad M_N(t) = e^{\lambda(e^t-1)}, \quad g_N(s) = \ln M_N(s) = \lambda(e^s - 1).$$

From the functions above, one may check that the first three raw moments are

$$E(N) = \lambda, \quad E(N^2) = M''(0) = \lambda + \lambda^2 \quad \mathrm{E}[N^3] = M'''(0) = \lambda + 3\lambda^2 + \lambda^3.$$

The $k-$th factorial moment is

$$\mathrm{E}[N(N-1)\dots(N-k+1)] = P^{(k)}(1) = \lambda^k, \quad k = 1, 2, \dots,$$

The second and third central moments are given by

$$Var(N) = M''(0) - (M'(0))^2 = \lambda, \quad E[(N-\mu_N)^3] = g'''(0) = \lambda.$$

Finally, the asymmetry coefficient is given by

$$\gamma_N = \frac{E[(N-\mu_N)^3]}{(Var(X))^{3/2}} = \frac{1}{\sqrt{\lambda}}.$$

Next, we present two useful properties of the Poisson distribution. The first result describes the additive property of the Poisson process.

Theorem 2.1. *Let $N_1, ..., N_n$ be independent Poisson random variables with parameters λ_1, λ_2, ..., λ_n. Then $N = N_1 + ... + N_n$ has a Poisson distribution with parameter $\lambda = \sum_{i=1}^{n} \lambda_i$.*

Proof. To prove this result, one may use the moment or the probability generating functions. Since $N_1, ..., N_n$ are independent random variables, we get that

$$M_N(t) = E(e^{tN}) = \prod_{i=1}^{n} E(e^{tN_i}) = \prod_{i=1}^{n} e^{\lambda_i(e^t-1)} = e^{\sum_{i=1}^{n}\lambda_i(e^t-1)}.$$

Therefore, N is a Poisson random variable with parameter $\sum_{i=1}^{n}\lambda_i$. ∎

Before we present the second property, we present an auxiliary result:

Theorem 2.2. *Suppose that the N is a Poisson with mean λ. Suppose that each event can be classified into one of m types with probabilities r_1, $r_2, ..., r_m$, (where $r_1 + r_2 + ... + r_m = 1$) independently of all the other events. Conditional to the event $\{N = n\}$, the joint probability function of N_1, ..., N_m is*

$$P(N_1 = n_1, N_2 = n_2, \cdots, N_m = n_m | N = n) = \frac{n!}{n_1! n_2! \cdots n_m!} r_1^{n_1} \cdots r_m^{n_m}.$$

The second property can be useful to model risks that can classified into different types. If, for any reason, the number of claims follows a Poisson distribution and each claim can be classified in type $1, 2, \cdots$. Then the number of claims type i is also a Poisson with a different parameter.

Theorem 2.3. *Suppose that the N is a Poisson with mean λ. Suppose that each event can be classified into one of m types with probabilities r_1, $r_2, ..., r_m$, (where $r_1 + r_2 + ... + r_m = 1$) independently of all the other events. Then the number of events N_1, ..., N_m classified in each type are independent Poisson random variables with means λr_1, $\lambda r_2, ..., \lambda r_m$.*

Proof. Taking into account the previous result:

$$P(N_1 = n_1, N_2 = n_2, \cdots, N_m = n_m | N = n) = \frac{n!}{n_1! n_2! \cdots n_m!} r_1^{n_1} \cdots r_m^{n_m}$$

Therefore,

$$\begin{aligned} P(N_1 = n_1, \cdots, N_m = n_m) &= P(N_1 = n_1, \cdots, N_m = n_m | N = n) P(N = n) \\ &= \frac{n!}{n_1! n_2! \cdots n_m!} r_1^{n_1} \cdots r_m^{n_m} \frac{e^{-\lambda}\lambda^n}{n!} \\ &= e^{-\lambda} \prod_{i=1}^{m} \frac{(r_i\lambda)^{n_i}}{n_i!} = \prod_{i=1}^{m} \frac{e^{\lambda r_i}(r_i\lambda)^{n_i}}{n_i!}. \end{aligned}$$

The last equality follows in light of the fact that $r_1 + \cdots r_m = 1$. Using the total probability law and letting $\mathbf{n} = (n_1, \cdots, n_{i-1}, n_{i+1}, \cdots, n_m)$, we get

$$
\begin{aligned}
P(N_i = n_i) &= \sum_{\mathbf{n}} P(N_1 = n_1, \cdots, N_m = n_m) \\
&= \frac{e^{-\lambda r_i}(r_i\lambda)^{n_i}}{n_i!} \underbrace{\sum_{\mathbf{n}} \frac{e^{-\lambda r_1}(r_1\lambda)^{n_1}}{n_1!} \cdots \frac{e^{-\lambda r_{i-1}}(r_{i-1}\lambda)^{n_{i-1}}}{n_{i-1}!} \frac{e^{-\lambda r_{i+1}}(r_{i+1}\lambda)^{n_{i+1}}}{n_{i+1}!} \cdots \frac{e^{-\lambda r_m}(r_m\lambda)^{n_m}}{n_m!}}_{=1} \\
&= \frac{e^{-r_i\lambda}(r_i\lambda)^{n_i}}{n_i!}.
\end{aligned}
$$

The third equality is obtained noticing that each term represents the probability function of a Poisson with a mean equal to $r_i\lambda$. The sum of each that terms is one because we are summing all the terms of a probability function. ∎

2.1.2 The Negative Binomial distribution

N is a negative binomial if its probability function is given by

$$p_k = \binom{r+k-1}{k} \left(\frac{\beta}{1+\beta}\right)^k \left(\frac{1}{1+\beta}\right)^r, \; k = 0, 1, ..., r > 0, \beta > 0$$

where

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

with $x > k-1$ in the last expression.

That random variable can be seen as a mixture of a Poisson where the structure distribution is a gamma, i.e. given $\Lambda = \lambda$, N is a Poisson random variable with mean λ, and λ is the observation of a random variable Λ gamma distributed with parameters $(\alpha = r, \theta = \beta)$.

$$
\begin{aligned}
p_k &= \Pr\{N = k\} = E[\Pr\{N = k|\Lambda\}] = \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!} \frac{\lambda^{r-1}e^{-\lambda/\beta}}{\beta^r\Gamma(r)} d\lambda \\
&= \frac{1}{k!\beta^r\Gamma(r)} \int_0^\infty e^{-\lambda/\left(\frac{\beta}{\beta+1}\right)} \lambda^{k+r-1} d\lambda \\
&= \binom{r+k-1}{k} \left(\frac{\beta}{1+\beta}\right)^k \left(\frac{1}{1+\beta}\right)^r, \; k = 0, 1, ...
\end{aligned}
$$

The probability, moment and cumulant generating functions are given by

$$P_N(z) = E[z^N] = E[E[z^N|\Lambda]] = E[e^{\Lambda(z-1)}] = (1 - \beta(z-1))^{-r}.$$

and

$$M_N(t) = (1-\beta(e^t-1))^{-r}, \quad \text{and} \quad g_N(t) = -r\ln(1-\beta(e^t-1)).$$

Finally, one may check that the expected value, the variance and third central moment are

$$E[N] = r\beta, \quad Var[N] = r\beta + r\beta^2, \quad E[(N-\mu_N)^3] = (r\beta + 3r\beta^2 + 2r\beta^3)$$

It is also possible to check that the Poisson distribution may be regarded as the limit of the negative binomial when $r \to \infty$, $\beta \to 0$, and the product $r\beta$ is constant ($=\lambda$).

Proposition 2.1. *Assume that N is random variable such that*

$$P_N(z) = (1-\beta(z-1))^{-r} \equiv P_N(z;r).$$

Further, consider that $r\beta = \lambda$. Then,

$$\lim_{r\to\infty} P_N(z;r) = \exp\left\{\lambda(z-1)\right\}.$$

Proof.

$$\begin{aligned}
\lim_{r\to\infty}\left(1-\frac{\lambda}{r}(z-1)\right)^{-r} &= \exp\left\{\lim_{r\to\infty} -r\ln\left(1-\frac{\lambda}{r}(z-1)\right)\right\} = \\
&= \exp\left\{-\lim_{r\to\infty}\frac{\ln\left(1-\frac{\lambda}{r}(z-1)\right)}{r^{-1}}\right\} = (\text{L´Hôpital´s rule}) \\
&= \exp\left\{\lim_{r\to\infty}\frac{\frac{\lambda}{r^2}(z-1)\left(1-\frac{\lambda}{r}(z-1)\right)^{-1}}{r^{-2}}\right\} = \\
&= \exp\left\{\lim_{r\to\infty}\frac{r\lambda(z-1)}{(r-\lambda(z-1))}\right\} = (\text{L´Hôpital´s rule}) \\
&= \exp\left\{\lambda(z-1)\right\}.
\end{aligned}$$

∎

The geometric distribution is a particular case of the negative binomial when $r = 1$. In this sense, the probability function is

$$p_k = \frac{1}{1+\beta}\left(\frac{\beta}{1+\beta}\right)^k, k = 0, 1, ...$$

For the geometric distribution, one can easily prove that

$$\Pr\{N > n\} = \sum_{k=n+1}^{\infty}\frac{1}{1+\beta}\left(\frac{\beta}{1+\beta}\right)^k = \frac{1}{1+\beta}\left(\frac{\beta}{1+\beta}\right)^{n+1}\frac{1}{1-\frac{1}{1+\beta}} = \left(\frac{\beta}{1+\beta}\right)^{n+1},$$

the second equality following in light of the properties of the geometric series. Additionally, it is easy to prove that the geometric distribution verifies memoryless, that is $\Pr\{N > m+n|N \geq m\} = \Pr\{N > n\}$.[1] Indeed, we can easily see that

$$\Pr\{N > m+n|N \geq m\} = \frac{\Pr\{N > m+n\}}{\Pr\{N \geq m\}} = \frac{\Pr\{N > m+n\}}{\Pr\{N > m-1\}} = \frac{\left(\frac{\beta}{1+\beta}\right)^{m+n+1}}{\left(\frac{\beta}{1+\beta}\right)^{m}}$$
$$= \left(\frac{\beta}{1+\beta}\right)^{n+1} = \Pr\{N > n\}$$

This is a property shared by a continuous random variable: the exponential. In the Loss Model's book the following interpretation is given: *Given that there are at least m claims, the probability distribution of the number of claims in excess of m does not depend on m.*

It is used to consider that the exponential divide the distributions in heavy tail and light tail: The negative binomial has a heavy tail when $r < 1$ and a light tail when $r > 1$.

2.1.3 The Binomial distribution:

N is a binomial random variable if its probability function is

$$p_k = \binom{m}{k} q^k (1-q)^{m-k}$$

It describes the situation were m independent risks are each subject to the probability q of making a claim. Additionally, the number of claims for each individual is a Bernoulli with parameter q, and the number of claims of the m independent and identical individuals is a Binomial (m, q).

The probability, moment and cumulant generating function are given by

$$P_N(z) = (1 + q(z-1))^m, \quad M_N(t) = (1 + q(e^t - 1))^m, \quad g_N(t) = m \ln(1 + q(e^t - 1)).$$

Additionally, if $p = 1 - q$, then

$$\mathrm{E}[N] = mq, \quad \mathrm{Var}[N] = mqp, \quad E[(N - \mu_N)^3] = mq - 3mq^2 + 2mq^3 = mqp(p - q).$$

We have already seen that the Poisson distribution can be obtained as a limit of the negative binomial distribution. A similar result can be stated with the binomial distribution.

[1] If we model the geometric distribution starting in 1, then the probability function is

$$p_k = \frac{1}{1+\beta}\left(\frac{\beta}{1+\beta}\right)^{k-1}, k = 1, 2, \cdots$$

The memoryless property for this case is $\Pr\{N > m+n|N > m\} = \Pr\{N > n\}$.

Proposition 2.2. *Assume that N is random variable such that*

$$P_N(z) = (1 + q(z-1))^m \equiv P_N(z;m).$$

Further, consider that $mq = \lambda$. Then,

$$\lim_{m\to\infty} P_N(z;m) = \exp\{\lambda(z-1)\}.$$

Proof.

$$\begin{aligned}
\lim_{m\to\infty}(1+q(z-1))^m &= \exp\left(\lim_{m\to\infty} m\ln(1+q(z-1))\right) \\
&= \exp\left(\lim_{m\to\infty} \ln(1+\frac{\lambda}{m}(z-1))\right) \\
&= \exp\left(\lim_{m\to\infty} \frac{\ln(1+\frac{\lambda}{m}(z-1))}{1/m}\right) \\
&= \exp\left(\lim_{m\to\infty} \frac{-(1+\frac{\lambda}{m}(z-1))^{-1}\frac{\lambda}{m^2}(z-1)}{-1/m^2}\right) \\
&= \exp\left(\lim_{m\to\infty} \frac{\lambda(z-1)}{1+\frac{\lambda}{m}(z-1)}\right) = \exp\{\lambda(z-1)\}
\end{aligned}$$

∎

2.1.4 The $(a, b, 0)$ class of distributions

The $(a, b, 0)$ class of distributions includes the Poisson, the negative binomial, and the binomial distributions.

Definition 2.1. *Let $p_k = \Pr\{N = k\}$, $k = 0, 1, 2\ldots$ be the probability function of N. Then, its distribution is a member of the $(a, b, 0)$ class of distributions if there exist constants a and b such that*

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

Table 1 sum up all the non-degenerated distributions belonging to the $(a, b, 0)$ class of distributions.

Example 2.1. *One may prove that the negative binomial distribution belongs to the $(a, b, 0)$ class of distributions noticing that*

$$\frac{\binom{r+k-1}{k}\left(\frac{1}{1+\beta}\right)^r\left(\frac{\beta}{1+\beta}\right)^k}{\binom{r+k-2}{k}\left(\frac{1}{1+\beta}\right)^r\left(\frac{\beta}{1+\beta}\right)^{k-1}} = \frac{\binom{r+k-1}{k}}{\binom{r+k-2}{k-1}}\left(\frac{\beta}{1+\beta}\right)$$

Table 1: $(a, b, 0)$ class of distributions

Distribution	Probability function	a	b
Poisson	$\dfrac{\mathrm{e}^{-\lambda}\lambda^k}{k!}$	0	λ
Negative Binomial	$\left(\begin{array}{c} r+k-1 \\ k \end{array} \right) \left(\frac{1}{1+\beta}\right)^r \left(\frac{\beta}{1+\beta}\right)^k$	$\frac{\beta}{1+\beta}$	$(r-1)\frac{\beta}{1+\beta}$
Binomial	$\left(\begin{array}{c} m \\ k \end{array} \right) q^k(1-q)^{m-k}$	$-\frac{q}{1-q}$	$(m+1)\frac{q}{1-q}$

where

$$\frac{\binom{r+k-1}{k}}{\binom{r+k-2}{k-1}} = \frac{\frac{\Gamma(r+k)}{\Gamma(k+1)\Gamma(r)}}{\frac{\Gamma(r+k-1)}{\Gamma(k)\Gamma(r)}} = \frac{\frac{(r+k-1)\Gamma(r+k-1)}{k\Gamma(k)}}{\frac{\Gamma(r+k-1)}{\Gamma(k)}} = \frac{r+k-1}{k} = \frac{r-1}{k} + 1.$$

This means that

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

and, therefore,

$$a = \frac{\beta}{1+\beta} \quad \text{and} \quad b = (r-1)\frac{\beta}{1+\beta}.$$

Example 2.2. *To verify that the Poisson distribution belongs to the $(a, b, 0)$ class of distributions, we only need to notice that*

$$\frac{p_k}{p_{k-1}} = \frac{\frac{e^{-\lambda}\lambda^k}{k!}}{\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}} = \frac{\lambda}{k}.$$

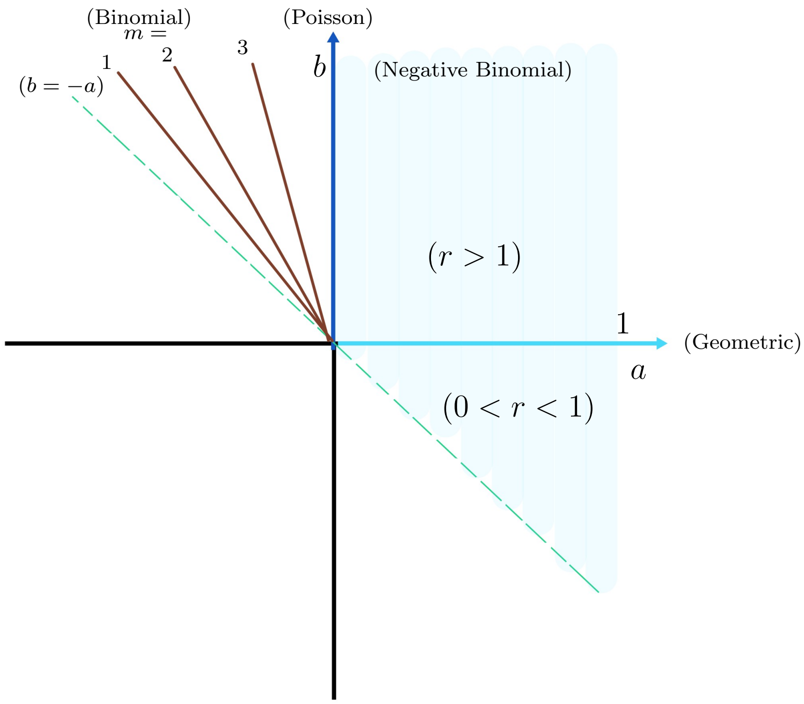
Therefore, $\frac{p_k}{p_{k-1}} = a + \frac{b}{k}$ with $a = 0$ and $b = \lambda$.

Exercise: Prove that the binomial distributions belong to the $(a, b, 0)$ class of distributions.

Theorem 2.4 shows that the Poisson, binomial, and binomial negative distributions are the only distributions in the $(a, b, 0)$ class. Before we prove Theorem 2.4, we present Figure 1 that depicts the class $(a, b, 0)$ of distributions in the space of parameters (a, b).

Theorem 2.4. *The Poisson, the negative binomial, and the binomial are the only distributions taking values on the non-negative integers that belong to the $(a, b, 0)$ class of distributions.*

Figure 1: $(a, b, 0)$ class of distributions



Proof. Firstly we can notice that Poisson, binomial and negative binomial distributions cover the following regions in the space of parameters (a, b):

- Poisson distribution

$$\{(a, b) : \ a = 0, b > 0\}$$

- Binomial distribution

$$\{(a, b) : \ a < 0, b = -(m + 1)a\}$$

- Negative binomial distribution

$$\{(a, b) : \ 0 < a < 1, b > -a\}.$$

Therefore, we prove the result showing that parameters (a, b) in the remaining regions do not provide non-degenerate distributions. The remaining regions are

1. $\{(a,b) : b \leq -a\}$
2. $\{(a,b) : a \geq 1, b > -a\}$
3. $\{(a,b) : a < 0, b > -a, b \neq -(m+1)a\}$.

In region 1., if $a + b = 0$, then we have a degenerate distribution. Otherwise, we have that $p_1 = (a+b)p_0 < 0$, which is not allowed. In region 2., we have that

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k} > a - \frac{a}{k} > \frac{k-1}{k}.$$

Therefore, for $k > 1$ we have

$$p_k > \frac{k-1}{k} p_{k-1} \Leftrightarrow p_k > \frac{k-1}{k} \times \frac{k-2}{k-1} \times \cdots \frac{1}{2} p_1 \Leftrightarrow p_k > \frac{1}{k} p_1.$$

If $\{p_k\}$ represents a probability function then $\sum_{k=1}^{\infty} p_k = 1 - p_0$. However, in situation 2. we have

$$\sum_{k=1}^{\infty} p_k > \sum_{k=1}^{\infty} \frac{1}{k} p_1 = p_1 \sum_{k=1}^{\infty} \frac{1}{k} = +\infty,$$

which is not possible.

To finalize this proof, we note that when $a < 0$ and $b > -a$, we have that $\lim_{k\to+\infty}\left(a + \frac{b}{k}\right) = a < 0$ which implies that $p_k \leq 0$ for values of k sufficiently large. We only have a probability distribution in this region, if there is a k such that $p_k = 0$ because this implies that $p_{k+n} = 0$, for all $n \in \mathbb{N}$. It is a matter of computations to check that there is a k such that $p_k = 0$ only when $b = -(m+1)a$. Indeed, one can see that, in this situation, we get $k = m + 1$. ∎

We finalize this section noticing that the members of the $(a, b, 0)$ family may be regarded as distributions that have as probability generating function, a function of the form

$$P_N(z; \theta) = B(\theta(z-1)),$$

where θ is a parameter and $B(.)$ is a function independent of θ. In the Poisson case, $\theta = \lambda$ and $B(x) = \mathrm{e}^x$; for the binomial $\theta = q$ and $B(x) = (1+x)^m$ and for the negative binomial $\theta = \beta$ and $B(x) = (1-x)^{-r}$.

2.1.5 Empirical analysis

It can be useful to recognize some characteristics of these distributions to choose which of the three better fit the data.

- The mean and variance of the Poisson are equal;

- The variance of the negative binomial is larger than the mean. In this case one may consider the negative binomial as an alternative to the Poisson;
- The variance of the binomial is smaller than its mean; It can be applied when the insurance company has m independent risk all with a probability q of making a claim;
- The binomial distribution has a finite support which may be useful to model some risks (The number of accidents per automobile).
- The relationship $\frac{p_k}{p_{k-1}} = a + \frac{b}{k}$ can be written as $k\frac{p_k}{p_{k-1}} = ak + b$. Therefore, if n_k represents the number of policies with k claims, the plot of $k\frac{n_k}{n_{k-1}}$ should be approximately a straight line. The shape of that line may indicate us the distribution that better fits the data.

Example 2.3. *Suppose that the claim frequency of 7263 motor insurance policies is the following*

number of occurrences k	0	1	2	3	4	5	6
n_k	6000	1000	200	50	10	3	0

Computing the ratio $k\frac{n_k}{n_{k-1}}$, we get the following

k	1	2	3	4	5	6
$k\frac{n_k}{n_{k-1}}$	0.1666667	0.4	0.75	0.8	1.5	0

Therefore, $k\frac{n_k}{n_{k-1}}$ is increasing in k which means that we can guess that the negative binomial fits better the data. Additionally, one can check that

$$mean = 0.2209831 \quad var = 0.293352,$$

which is coherent with the comments above.

2.1.6 R statistical software

In this section, we introduce some useful functions in R that will allow us to compute quickly some probabilities. The "actuar" package has most of the functions we need in Risk Theory. Although we do not need it in this section, we can start with

```
install.packages("actuar")

library(actuar)
```

In the package “stats” we can find the following functions, that represent respectively the probability function, the distribution function and the quantile function of the binomial distribution.

```
dbinom(x, size, prob, log = FALSE)
pbinom(q, size, prob, lower.tail = TRUE, log.p = FALSE)
qbinom(p, size, prob, lower.tail = TRUE, log.p = FALSE)
```

Similar functions can be found for Poisson and negative binomial distributions.

```
dpois(x, lambda, log = FALSE)
ppois(q, lambda, lower.tail = TRUE, log.p = FALSE)
qpois(p, lambda, lower.tail = TRUE, log.p = FALSE)
```

and

```
dnbinom(x, size, prob, mu, log = FALSE)
pnbinom(q, size, prob, mu, lower.tail = TRUE, log.p = FALSE)
qnbinom(p, size, prob, mu, lower.tail = TRUE, log.p = FALSE)
rnbinom(n, size, prob, mu)
```

We note that the parameterization of the negative binomial used in these functions is different from the one adopted in this course. You have to notice that $prob = \frac{1}{1+\beta}$. An alternative way is fixing $mu = r\beta$. The geometric distribution can be accessed through the functions

```
dgeom, pgeom, qgeom
```

or by choosing the size of the negative binomial distribution equal to 1.

2.2 The Poisson process

The number of claims over time associated to a risk is modeled by a stochastic process, more concretely, by a counting process.

Definition 2.2. *A stochastic process N_t, $t \geq 0$ is a collection of random variables, indexed by the variable t (which often represents time).*

Definition 2.3. *A counting process is a stochastic process $\{N(t) : t \geq 0\}$ such that $N(t)$ is a non-negative integer and for any $t \geq s$, $N(t) \geq N(s)$.*

The Poisson stochastic process is used in many different applications and it is a particular case of counting processes. It can be defined in different, but equivalent ways.

Definition 2.4. *A counting process $\{N(t) : t \geq 0\}$, with $N(0) = 0$, is an homogeneous Poisson process, or just a Poisson process, with intensity λ, if it satisfies the following postulates:*

1) $\{N(t) : t \geq 0\}$ has independent and stationary increments;

2) The random variable $N(t)$ follows a Poisson distribution with mean λt

Taking into account that $N(t)$ follows a Poisson distribution, then one may notice that, for a small h

$$Pr(N(h) = 0) = e^{-\lambda h} = 1 - \lambda h + \sum_{i=2}^{\infty} \frac{(-\lambda h)^i}{i!} = 1 - \lambda h + o(h).$$[2]

Additionally,

$$Pr(N(h) = 1) = e^{-\lambda h}\lambda h = \lambda h - \sum_{i=2}^{\infty} \frac{(-\lambda h)^i}{i!} = \lambda h + o(h).$$

We can conclude that $Pr(N(h) \geq 2) = o(h)$. This motivates the equivalent and alternative definition.

Definition 2.5. *A counting process $\{N(t) : t \geq 0\}$, with $N(0) = 0$, is an homogeneous Poisson process, or just a Poisson process, with intensity λ, if it satisfies the following postulates:*

1) $\{N(t) : t \geq 0\}$ has independent and stationary increments;

2) $Pr(N(h) = 1) = \lambda h + o(h)$, and $Pr(N(t) \geq 2) = o(h)$.

Discussing of the postulates:

- Independent increments: Excludes chain reactions. A fire can originate another fire. This difficulty may be, sometimes, overtaken redefining the risk unit. This is the case of fire insurance. But not the case for contagious diseases or epidemics.

- Stationary increments and $Pr(N(t) = 1) = \lambda h + o(h)$: There are situations where they are not verified. An example is when there are seasonality involved. In some cases time may be divided into sub-intervals, to obtain sub-processes with different intensities.

  If we are only interested in the number of claims over a finite time interval the Poisson distribution remains valid, even when there is a deterministic tendency on the claim frequency.

- $Pr(N(t) \geq 2) = o(h)$: This difficulty may be overtaken. For example an accident involving two cars, insured in the same company, and when both drivers are considered responsible, may be considered as just one claim.

[2] A function f is an infinitesimal with h and is denoted $o(h)$ when $\lim_{h \to 0^+} \frac{f(h)}{h} = 0$.

Properties of the Poisson process: Since we know that the random variable $N(t)$ follows a Poisson distribution, then, as previously stated,

$$P_{N(t)}(z) = e^{\lambda t(z-1)}, \quad M_{N(t)}(r) = e^{\lambda t(e^r-1)}, \quad g_{N(t)}(s) = \ln M_N(r) = \lambda t(e^r - 1).$$

From the functions above, one may check that the first three raw moments are

$$E(N(t)) = \lambda t, \quad Var(N(t)) = \lambda t, \quad \text{and} \quad \gamma_{N(t)} = \frac{1}{\sqrt{\lambda t}}.$$

The covariance between the stochastic process in two different instants of time $0 < s < t$ is given by

$$\begin{aligned} cov(N(t), N(s)) &= Cov(N(t) - N(s) + N(s), N(s)) \\ &= Cov(N(t) - N(s), N(s)) + Var(N(s)) \\ &= Var(N(s)) = \lambda s. \end{aligned}$$

In general, for any $s, t > 0$ one may write

$$cov(N(t), N(s)) = \lambda \min(s, t).$$

There are some distributions related to the Poisson process as we can see in the next results. Let W_k be the time of the k^{th} event, with $k = 1, 2, \cdots$. The difference $T_k = W_{k+1} - W_k$ represents the time between events k and $k + 1$. the variables T_k represent the times of performance of the process in state k.

Proposition 2.3. *The interarrival times T_k in a Poisson process are i.i.d random variables exponential distributed with mean $1/\lambda$.*

Proof. The proof is straightforward:

$$\begin{aligned} Pr(T_k > t | W_k = s) &= P(N(W_k + t) = k, N(W_k) = k) = P(N(W_k + t) - N(W_k) = 0) \\ &= P(N(t) = 0) = e^{-\lambda t} \end{aligned}$$

Therefore, T_k is an exponential distribution with mean $1/\lambda$. Additionally,

$$\begin{aligned} Pr(T_k > t, T_{k+1} > z) &= Pr(T_k > t, T_{k+1} > z | W_k = s) \\ &= Pr(T_{k+1} > z | T_k > t, W_k = s) Pr(T_k > t | W_k = s) \\ &= Pr(T_{k+1} > z | W_{k+1} = s + t) Pr(T_k > t | W_k = s) \\ &= Pr(T_{k+1} > z) Pr(T_k > t). \end{aligned}$$

∎

The Poisson process is related to the binomial distribution in two different ways.

Proposition 2.4. *Let $u < t$ and $k \leq n$ then the distribution of $N(u)$ conditional to the information that $N(t) = n$ is $Bin(n, u/t)$*

Proof.

$$
\begin{aligned}
Pr(N(u) = k | N(t) = n) &= \frac{Pr(N(u) = k, N(t) = n)}{P(N(t) = n)} = \frac{Pr(N(u) = k, N(t) - N(u) = n - k)}{P(N(t) = n)} \\
&= \frac{\frac{e^{-\lambda u}(\lambda u)^k}{k!} \frac{-e^{\lambda(t-u)}(\lambda(t-u))^{n-k}}{(n-k)!}}{\frac{e^{-\lambda t}(\lambda t)^n}{n!}} = \frac{u^k (t-u)^{n-k}}{t^n} \times \frac{n!}{k!(n-k)!} \\
&= \binom{n}{k} (u/t)^k (1 - u/t)^{n-k}
\end{aligned}
$$

∎

Proposition 2.5. *Let $\{N_1(t) : t \geq 0\}$ and $\{N_2(t) : t \geq 0\}$ be two independent Poisson processes with intensities λ_1 and λ_2. Then, the distribution of $N_1(t)$ conditional to the information that $N_1(t) + N_2(t) = n$ is $Bin\left(n, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$*

Proof. The proof is similar to the previous one. ∎

Another distribution related to the Poisson process is the uniform.

Proposition 2.6. *If $s < t$, then the distribution of W_1 conditional to the information that $N(t) = 1$ is uniform in the interval $(0, t)$.*

Proof.

$$
\begin{aligned}
P(W_1 \leq s | N(t) = 1) &= \frac{P(N(s) = 1, N(t) - N(s) = 0)}{P(N(t) = 1)} = \frac{P(N(s) = 1) P(N(t) - N(s) = 0)}{P(N(t) = 1)} \\
&= \frac{e^{-\lambda(t-s)} e^{-\lambda s} \lambda s}{e^{-\lambda t} \lambda t} = \frac{s}{t}.
\end{aligned}
$$

∎

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

In some situations, the $(a, b, 0)$ class of distributions is not adequate to model insurance data. Therefore, we generalize that class.

Definition 2.6. *Let $p_k = \Pr\{N = k\}$, $k = 0, 1, 2, \ldots$ be the p.f. of a discrete r.v. taking values at the nonnegative integers. This distribution belongs to the $(a, b, 1)$ class of distributions if there are constants a and b such that*

$$
p_k = \left(a + \frac{b}{k}\right) p_{k-1}, \quad k = 2, 3 \ldots .
$$

We note that the class $(a, b, 1)$ of distributions generalize the class $(a, b, 0)$ of distributions because the recursion presented starts in p_2 and not in p_1. Therefore, we may obtain distributions from the $(a, b, 0)$ class by setting $P(N = 0) = 0$ or modifying the probability at 0, $P(N = 0) = p_0^M$.

2.3.1 Zero-modified and zero-truncated distributions

If we change the probability of having zero claims in the distributions of the $(a, b, 0)$ class, then we have to re-scale all the distribution to ensure that the relationship in Definition 2.6 is still satisfied.Assume that, we assign a new probability at 0, p_0^M, then we nee to find c such that

$$p_0^M + c\underbrace{\sum_{k=1}^{\infty} p_k}_{1-p_0} = 1 \Leftrightarrow c = \frac{1 - p_0^M}{1 - p_0}.$$

Therefore, if we modify the probability at zero of any member of the $(a, b, 0)$ class, the modified probabilities are

$$p_k^M = \frac{1 - p_0^M}{1 - p_0} p_k, \ k = 1, 2, \ldots,$$

The modified members of the $(a, b, 0)$ class of distributions are the *zero-modified Poisson* , the *zero-modified binomial* and the *zero-modified negative binomial*. Note that these distributions can be regarded as a mixture of a member of the class $(a, b, 0)$ with a degenerate distribution at the origin. When $p_0^M = 0$ the modified distribution is called *truncated at zero*.

The generating probability function can be obtained as follows

$$\begin{aligned} P_N^M(z) &= \sum_{k=0}^{\infty} p_k^M z^k = p_0^M - p_0 \frac{1 - p_0^M}{1 - p_0} + \frac{1 - p_0^M}{1 - p_0} \sum_{k=0}^{\infty} p_k z^k \\ &= \frac{p_0^M - p_0}{1 - p_0} + \frac{1 - p_0^M}{1 - p_0} P_N(z). \end{aligned}$$

When $p_0^M = 0$, then

$$P_N^M(z) = \frac{P_N(z) - p_0}{1 - p_0}.$$

2.3.2 The extended modified negative binomial

As we have seen in the $(a, b, 0)$ class of distribution, the negative binomial is defined in the space of parameters $\{(a, b) \in \mathbb{R}^2 : 0 < a < 1, a + b > 0\}$. The restriction $a + b > 0$, comes from the fact that $p_1 = (a + b)p_0$. For the $(a, b, 1)$ class of distributions, we have that $a + b/k > 0$, $\forall k \geq 2$ if and only if $a + b/2 > 0$, which is equivalent to have $r > -1$.

Therefore, space parameter of the negative binomial can be extended to case where $r > -1$, $r \neq 0$. For the negative binomial the relationship

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

with parameters $a = \frac{\beta}{1+\beta}$ and $b = (r-1)\frac{\beta}{1+\beta}$. The extended modified negative binomial verifies the same relationship with the same parameters

$$\frac{p_k^M}{p_{k-1}^M} = a + \frac{b}{k}, \quad \text{for} \quad k = 2, 3, \dots .$$

Taking into account that $a + b/k > 0$, $\forall k \geq 2$ and p_0^M is fixed and $0 < p_0^M < 1$, the distribution is well defined if there is $0 < p_1^M < 1$ such that

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

In light of the previous computations, we have

$$p_1^M = \frac{1 - p_0^M}{1 - \left(\frac{1}{1+\beta}\right)^r} r \left(\frac{\beta}{1+\beta}\right) \left(\frac{1}{1+\beta}\right)^r,$$

that is positive for $\beta > 0$ and $0 < p_0^M < 1$. To prove that $\sum_{k=0}^{\infty} p_k^M = 1$, we can check that

$$\sum_{k=1}^{\infty} p_k^M = \frac{1 - p_0^M}{1 - \left(\frac{1}{1+\beta}\right)^r} \sum_{k=1}^{\infty} p_k = \frac{1 - p_0^M}{1 - \left(\frac{1}{1+\beta}\right)^r} \left(1 - \left(\frac{1}{1+\beta}\right)^r\right) = 1 - p_0^M$$

One interesting case is the truncated extended negative binomial for $r > -1, \neq 0$, where $p_0^M = 0$. In fact for that case, we get

$$p_1^T = \frac{1}{1 - \left(\frac{1}{1+\beta}\right)^r} r \left(\frac{\beta}{1+\beta}\right) \left(\frac{1}{1+\beta}\right)^r$$

Finally, one can prove that the limiting case of the truncated extended negative binomial when $r \to 0$ is the **logarithmic distribution**. To deduce the distribution for that limiting case, we note that

$$a = \frac{\beta}{1+\beta} \quad \text{and} \quad b = -\frac{\beta}{1+\beta}.$$

Therefore,

$$\begin{aligned} p_k^T &= \left(a + \frac{b}{k}\right) p_{k-1}^T = \\ &= \left(\frac{\beta}{1+\beta} - \frac{\frac{\beta}{1+\beta}}{k}\right) p_{k-1}^T = \frac{\beta}{1+\beta} \frac{k-1}{k} p_{k-1}^T. \end{aligned}$$

Hence

$$
\begin{aligned}
p_k^T &= \left(\frac{\beta}{1+\beta}\right)^{k-1} \frac{k-1}{k} \times \frac{k-2}{k-1} \times \cdots \times \frac{1}{2} p_1^T = \\
&= p_1^T \left(\frac{\beta}{1+\beta}\right)^{k-1} \frac{1}{k}.
\end{aligned}
$$

Since $\sum_{i=1}^{\infty} p_k^T = 1$ and taking into account that $ln(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k}$, $0 < x < 1$, we get that

$$
\begin{aligned}
p_1^T &= \frac{1}{\left(\frac{\beta}{1+\beta}\right)^{-1} \sum_{k=0}^{\infty} \left(\frac{\beta}{1+\beta}\right)^k \frac{1}{k}} = \\
&= \frac{\beta}{1+\beta} \frac{1}{-\ln\left(1 - \frac{\beta}{1+\beta}\right)} = \\
&= \frac{\beta}{1+\beta} \frac{1}{-\ln\left(\frac{1}{1+\beta}\right)} = \frac{\beta}{(1+\beta)\ln(1+\beta)}.
\end{aligned}
$$

Replacing the of p_1^T in the expression of p_k^T

$$
p_k^T = \left(\frac{\beta}{1+\beta}\right)^k \frac{1}{k\ln(1+\beta)}, \ k = 1, 2, ...
$$

The logarithmic distribution is a member of the $(a, b, 1)$ class of distributions with the following parameters:

$$
p_0 = 0, \quad a = \frac{\beta}{1+\beta}, \quad b = \frac{-\beta}{1+\beta}, \quad \text{with } \beta > 0.
$$

2.3.3 R statistical software

The distributions belonging to the $(a, b, 1)$ class are represented in the "actuar" package. Similarly to the Poisson, binomial and negative binomial, we have the probability function, distribution function and quantile function for the zero truncated and zero modified distributions.

- Zero modified negative binomial:

```
dzmnbinom(x, size, prob, p0, log = FALSE)
pzmnbinom(q, size, prob, p0, lower.tail = TRUE, log.p = FALSE)
qzmnbinom(p, size, prob, p0, lower.tail = TRUE, log.p = FALSE)
```

The main different between these functions and the ones presented for the negative binomial case is the argument $p0$, which allows us to modify the probability of the negative binomial at 0. Although there are specific functions to the zero truncated negative binomial at the "actuar" package

```
dztnbinom; pztnbinom; qztmnbinom
```

we may access these functions fixing $p0 = 0$ in the zero modified negative binomial. We should notice that the parameter size must be positive and the parameterization through the mean is not allowed. Therefore, in case we have $-1 < r < 0$ (extended modified/truncated negative binomial), we have to define our own function. For instance,

```
detnbinom <- function(v, r, beta) {
    if(r<=-1){print("NAA")
    } else {ifelse(v==0,0,
        gamma(r+v)/(gamma(r)*factorial(v))*
        (beta/(1+beta))^v*(1/(((1+beta)^r)-1)))
    }}
```

2.4 Compound frequency models

Let N be a counting distribution with probability generating function $P_N(z)$ and let $\{M_i\}$ be a sequence of i.i.d. counting random variables, independent from N, with probability generating function $P_M(z)$. The probability generating function of

$$S = M_0 + M_1 + M_2 + \ldots + M_N,$$

with $M_0 \equiv 0$, is

$$P_S(z) = P_N(P_M(z)).$$

A possible interpretation consists on considering N the number of accidents and M_i the number of claims from accident i. S would represent the total number of claims.

Example 2.4. Let N and $\{M_i\}$ have Poisson distribution, with parameters λ_1 and λ_2 respectively. The probability generating function of S is

$$P_S(z) = \ e^{\lambda_1(\ e^{\lambda_2(z-1)}-1)}$$

and it is called Poisson-Poisson or Neyman Type A.

If N is a Poisson with parameter λ and $\{M_i\}$ is a general counting distribution, then the probability generating function of S is given by

$$P_N(P_M(z)) = e^{\lambda(P_M(z)-1)}.$$

Example 2.5. *The compound geometric distribution with a secondary geometric distribution does not create a new distribution. Indeed, the geometric-geometric distribution is a special case of a zero-modified geometric. To see this, note that the probability generating function of a zero modified geometric with parameter β, P_{ZMG} is given by*

$$
\begin{aligned}
P_{ZMG}(z) &= \frac{p_0^M - (1+\beta)^{-1}}{1-(1+\beta)^{-1}} + \frac{1-p_0^M}{1-(1+\beta)^{-1}}[1-\beta(z-1)] \\
&= \frac{1+\beta}{\beta} \times \frac{\left(p_0^M - (1+\beta)^{-1}\right)(1-\beta(z-1)) + 1 - p_0^M}{1-\beta(z-1)} \\
&= \frac{1-(p_0^M(1+\beta)-1)(z-1)}{1-\beta(z-1)}.
\end{aligned}
$$

Consider now a compound distribution where the primary distribution is a geometric with β_1 and a secondary geometric distribution with parameter β_2. The probability generating function of the compound model will be

$$P_S(z) = \left[1-\beta_1\left([1-\beta_2(z-1)]^{-1}-1\right)\right]^{-1}.$$

Taking into account that

$$[1-\beta_2(z-1)]^{-1} - 1 = \frac{\beta_2(z-1)}{1-\beta_2(z-1)},$$

we get that

$$P_S(z) = \left[1-\beta_1 \times \frac{\beta_2(z-1)}{1-\beta_2(z-1)}\right]^{-1} = \frac{1-\beta_2(z-1)}{1-\beta_2(1+\beta_1)(z-1)}.$$

Therefore, we get the probability generating function of a zero modified geometric fixing

$$\beta_2(1+\beta_1) = \beta \quad \textit{and} \quad \beta_2 = p_0^M(1+\beta) - 1.$$

Let $p_n = \Pr\{N = n\}$, $n = 0, 1, 2, \ldots$, $f_n = \Pr\{M_i = n\}$, $n = 0, 1, 2, \ldots$ and $g_n = \Pr\{S = n\}$, $n = 0, 1, 2, \ldots$. Then

$$
\begin{aligned}
g_k &= \Pr\{S = k\} = \sum_{n=0}^{\infty} \Pr\{M_0 + M_1 + \ldots + M_n = k | N = n\} \Pr\{N = n\} \\
&= \sum_{n=0}^{\infty} \Pr\{M_0 + M_1 + \ldots + M_n = k\} \Pr\{N = n\} \\
&= \sum_{n=0}^{\infty} p_n f_k^{*n}, \qquad\qquad i = 0, 1, 2, \ldots,
\end{aligned}
$$

where f^{*n} is the n-fold convolution of f. An explanation about convolutions is provided in the third appendix.

Example 2.6. *Consider a compound frequency model with primary distribution N characterized by $P(N=n)=1/3$, for $n=0,1,2$, and secondary distribution M characterized by*

$$P(M=n)=\begin{cases}1/3, & n=1\\ 2/3, & n=2\end{cases}.$$

The distribution of S can be computed using convolutions as follows

k	f_n^{*0}	f_n^{*1}	f_n^{*2}	$f_S(k)$	$F_S(k)$
0	1			1/3	1/3
1		1/3		1/9	4/9
2		2/3	1/9	7/27	19/27
3			4/9	4/27	23/27
4			4/9	4/27	1
$P(N=n)$	1/3	1/3	1/3		

Next, we present the Panjer recursion formula.

Theorem 2.5. *Assume that the primary distribution N is a member of the $(a,b,0)$ class, then*

$$\begin{aligned} g_k &= \frac{1}{1-af_0}\sum_{j=1}^{k}\left(a+b\frac{j}{k}\right)f_j g_{k-j}, \quad k=1,2,\ldots \\ g_0 &= P_N(f_0). \end{aligned}$$

Proof. It is known that $g_0=P_S(0)=P_N(P_M(0))=P_N(f_0)$.

We start by noticing that

$$E\left[a+\frac{bM_1}{k}\middle|M_1+\cdots+M_j=k\right]=a+\frac{b}{j}.$$

Additionally, we can check that

$$E\left[a+\frac{bM_1}{s}\middle|M_1+\cdots+M_j=k\right]=\sum_{h=0}^{k}\left(a+\frac{bh}{k}\right)P\left(M_1=h\middle|\sum_{i=1}^{j}M_i=k\right)$$
$$=\sum_{h=0}^{k}\left(a+\frac{bh}{k}\right)\frac{P(M_1=h)P\left(\left(\sum_{i=1}^{j}M_i\right)-M_1=k-h\right)}{P\left(\sum_{i=1}^{j}M_i=k\right)}.$$

Taking into account the convolution formula, we can see that, for $k = 1, 2, \cdots$

$$\begin{aligned}
g_k &= \sum_{j=1}^{\infty} P\left(\sum_{i=1}^{j} M_i = k\right) p_j = \sum_{j=1}^{\infty} P\left(\sum_{i=1}^{j} M_i = k\right)\left(a + \frac{b}{j}\right) p_{j-1} \\
&= \sum_{j=1}^{\infty}\sum_{h=0}^{k}\left(a + \frac{bh}{k}\right) P(M_1 = h) P\left(\left(\sum_{i=1}^{j} M_i\right) - M_1 = k - h\right) p_{j-1} \\
&= \sum_{h=0}^{k}\left(a + \frac{bh}{k}\right) P(M_1 = h) \sum_{j=1}^{\infty} P\left(\left(\sum_{i=1}^{j} M_i\right) - M_1 = k - h\right) p_{j-1} \\
&= \sum_{h=0}^{k}\left(a + \frac{bh}{k}\right) P(M_1 = h) P(S = k - h) = \sum_{j=0}^{k}\left(a + \frac{bj}{k}\right) f_j g_{k-j}
\end{aligned}$$

∎

From the previous result, one may notice that if the primary distribution is a Poisson random variable, then the recursion can be simplified.

Lemma 2.1. *Assume that the primary distribution N is a Poisson, then*

$$\begin{aligned}
g_k &= \frac{\lambda}{k}\sum_{j=1}^{k} j f_j g_{k-j}, \qquad k = 1, 2, \ldots \\
g_0 &= \exp(-\lambda(1 - f_0)).
\end{aligned}$$

Example 2.7. *Let N be an binomial distribution with parameters $m = 2$ and $q = 0.4$, and M a discrete random variable with probability function*

$$f_n = \begin{cases} 1/3, & n = 0 \\ 2/3, & n = 1 \\ 0, & \textit{otherwise} \end{cases}$$

If S is the compound frequency random variable, its probability function, g_n, can be computed through the Panjer recursion formula. Taking into account that $a = -2/3$ and $b = 2$, then

$$g_k = \begin{cases} (1 + 0.4(1/3 - 1))^2, & k = 0 \\ \frac{1}{1+2/9}\left(-2/3 + 2\right) 2/3 \times g_0, & k = 1 \\ \frac{1}{1+2/9}\left(-2/3 + 2 \times 1/2\right) 2/3 g_1, & k = 2 \end{cases}. = \begin{cases} 0.53778, & k = 0 \\ 0.39111, & k = 1 \\ 0.07111, & k = 2 \end{cases}$$

If the primary distribution is a member of the $(a, b, 1)$ class, then the previous recursion formula has to be adjusted.

Theorem 2.6. *For the model here described and when N is a member of the $(a,b,1)$ family,*

$$\begin{aligned} g_k &= \frac{(p_1-(a+b)p_0)f_k+\sum_{j=1}^{k}(a+bj/k)\,f_j g_{k-j}}{1-af_0},\ k=1,2,\ldots \\ g_0 &= P_N(f_0). \end{aligned}$$

Next result shows us that the sum of compound Poisson processes is still a compound Poisson process

Theorem 2.7. *Suppose that S_i has a compound Poisson distribution with Poisson parameter λ_i and secondary distribution $\{q_n^i : n = 0,1,2,\ldots\}$, for $i = 1,2,3,\ldots,k$. Suppose also that $S_1, S_2,\ldots,S_k$ are independent random variables. Then $S = S_1+\ldots+S_k$ also has a compound Poisson distribution with parameter $\lambda = \lambda_1+\ldots+\lambda_k$ and secondary distribution $\{q_n : n = 0,1,2,\ldots\}$ where $q_n = [\lambda_1 q_{1,n}+\ldots+\lambda_n q_{k,n}]/\lambda$.*

Proof. Assuming that $Q_i(z)$ represents the probability generating function of the secondary distribution i, then pgf of S_i is given by

$$P_{S_i} = e^{\lambda_i(Q_i(z)-1)}.$$

Taking into account that

$$\begin{aligned} P_S(z) &= E\left(z^{\sum_{i=1}^{k} S_i}\right) = \prod_{i=1}^{k} P_{S_i} = \prod_{i=1}^{k} e^{\lambda_i(Q_i(z)-1)} \\ &= e^{\sum_{i=1}^{k}\lambda_i(Q_i(z))-1)} = e^{\sum_{i=1}^{k}\lambda_i(Q_i(z)))-\lambda} \\ &= e^{\lambda\left(\sum_{i=1}^{k}\frac{\lambda_i}{\lambda}(Q_i(z)))-1\right)}. \end{aligned}$$

∎

Example 2.8. *We have already seen that when N_1 and N_2 are two independent Poisson random variables with parameters λ_1 and λ_2, then $N_1+N_2 \sim Poi(\lambda)$, with $\lambda = \lambda_1+\lambda_1$. The result fails if we consider nN_1+mN_2 where $n \neq m \in \mathbb{N}$. Noticing that*

$$nN_1+mN_2 = n\sum_{i=0}^{N_1} M_{1,i} + m\sum_{i=0}^{N_2} M_{2,i}$$

with $M_{1,0} = M_{2,0} = 0$ and $M_{1,i} = M_{2,i} = 1$, for all $i = 1,2,\cdots$, we can guess that nN_1+mN_2 is a compound Poisson process with parameter $\lambda = \lambda_1+\lambda_2$ and secondary distribution

$$f_M(x) = \begin{cases} \frac{\lambda_1}{\lambda}, & x = n \\ \frac{\lambda_2}{\lambda}, & x = m \end{cases}.$$

To prove this result, one can compute the probability generating function

$$P_{nN_1+mN_2}(z) = E(z^{nN_1+mN_2}) = E\left((z^n)^{N_1}\right) E\left((z^m)^{N_2}\right)$$
$$= e^{\lambda_1(z^n-1)}e^{\lambda_2(z^m-1)} = e^{\lambda\left(\frac{\lambda_1}{\lambda}z^n+\frac{\lambda_2}{\lambda}z^m-1\right)}.$$

As we have seen previously the members of the $(a, b, 0)$ family may be regarded as distributions that have as probability generating function, a function of the form

$$P_N(z; \theta) = B(\theta(z-1)), \tag{1}$$

where θ is a parameter and $B(.)$ is a function independent of θ. Next result shows that changing the probability at the origin in the secondary distribution does not create a new compound distribution (only changes the parameter).

Theorem 2.8. *If $P_N[z;\theta] = B(\theta(z-1))$ for given θ and $B(z)$ independent of θ, then $P_S(z) = P_N[P_M(z);\theta]$ can be written as*

$$P_S(z) = P_N[P_M^T(z); \theta(1-f_0)],$$

where $P_M^T(z)$ is the p.g.f. of the secondary distribution truncated at the origin.

Proof.

$$P_S(z) = P_N[P_M(z);\theta] = B(\theta(P_M(z)-1))$$

Taking into account that

$$P_M^T(z)(1-f_0) = P_M(z) - f_0 \Leftrightarrow P_M(z) - 1 = (P_M^T(z)-1)(1-f_0)$$

we get

$$P_S(z) = P_N[P_M(z);\theta] = B(\theta(1-f_0)(P_M^T(z)-1)) = P_N[P_M^T(z);\theta(1-f_0)].$$

∎

Example 2.9. *Consider a Poisson-Poisson distribution with parameters λ_1 and λ_2. The probability generating function is*

$$P_S(z) = e^{\lambda_1(e^{\lambda_2(z-1)}-1)} = B(\lambda_1(P_M(z)-1)),$$

where $B(z) = e^z$. Taking into account that

$$(P_M^T(z)-1)(1-e^{-\lambda_2}) = \left(\frac{e^{\lambda_2(z-1)} - e^{-\lambda_2}}{1-e^{-\lambda_2}} - 1\right)(1-e^{-\lambda_2}) = e^{\lambda_2(z-1)} - e^{-\lambda_2} - 1 + e^{-\lambda_2}$$
$$= e^{\lambda_2(z-1)} - 1 = P_M(z) - 1.$$

Therefore,

$$P_S(z) = B(\lambda_1(P_M(z)-1)) = B(\lambda_1(1-e^{-\lambda_2})(P_T^M(z)-1)),$$

which means that the Poisson-Poisson distribution can be obtained as a compound Poisson where secondary distribution is a zero-truncated Poisson. The parameter of the Poisson is $\lambda_1(1-e^{-\lambda_2})$ and the parameter of the zero-truncated Poisson is λ_2.

2.4.1 R statistical software

R provides a function that allows us to compute probabilities when the frequency is modeled by a compound distribution.

```
aggregateDist(method = c("recursive", "convolution", "normal",
"npower", "simulation"), model.freq = NULL, model.sev = NULL,
p0 = NULL, x.scale = 1, convolve = 0, moments, nb.simul, ...,tol = 1e-06,
maxit = 500, echo = FALSE)
```

For now, we have only seen two methods: the recursive one (Panjer's recursion) and the convolutions. In R documentation, you can read a full explanation of this function. One can compute the probabilities in Example 2.7 with the following code:

```
library("actuar")
fx <- c(1/3,2/3)
Fs <- aggregateDist(method = "recursive", model.freq = "binomial",
model.sev = fx, size=2, prob=0.4, x.scale = 1)
diff(Fs)[1:3]
```

2.5 Mixed frequency distributions

A natural way to extend counting distributions is assuming that some of the parameters are themselves random variables. If $N|\Theta = \theta$ is a random variable with pgf $P(z;\theta)$, Θ is a random variable with probability/density function u and $p_k = P(N = k)$, then we get that

- if Θ is a discrete random variable we have a discrete mixture distribution and

$$p_k = \sum P(N = k|\Theta = \theta)u(\theta).$$

- if Θ is a continuous random variable we have a continuous mixture distribution and

$$p_k = \int p_k(\theta)u(\theta)d\theta$$

The probability generating function of N is

$$\begin{aligned} P_N(z) &= E\left(z^N\right) = E\left(E\left(z^N|\Theta\right)\right) = E\left(P_N(z;\Theta)\right) \\ &= \int P_{N|\Theta=\theta}(z)dU(\theta), \end{aligned}$$

where U is the distribution function of Θ. Often, Θ is known as risk parameter and U as structure distribution.

2.5.1 Discrete mixture distributions

We may have a discrete mixture of random variables when the total number of claims can be divided into different types. For instance, in health insurance a claim may be from a medical appointment, a surgery, an urgency, different treatments, dental care, etc.

Definition 2.7. *A random variable N is a $k-$point mixture of $N_1, N_2, \cdots, N_k$ if*

$$F_N(i) = a_1 F_{N_1}(i) + \cdots + a_k F_{N_k}(i)$$

where the parameters $a_i \leq 0$ for $i = 1, \cdots, k$ and $\sum_{i=1}^{k} a_1 = 1$

As previously introduced, one may define a discrete random variable Θ that takes values in $\{1, \cdots, k\}$ with probability function $P(\Theta = i) = a_i$, for $i = 1, \cdots, k$. In this case, $X_\Theta = X|\Theta$ and

$$\begin{aligned} P(N = i) &= \sum_{j=1}^{k} P(N = i|\Theta = j)P(\Theta = j) \\ &= \sum_{j=1}^{k} P(N_j = i)a_j \end{aligned}$$

and the probability generating function

$$P_N(z) = E(z^N) = E(E(z^N|\Theta)) = \sum_{j=1}^{k} a^j E(z^{N_j}) = \sum_{j=1}^{k} a_j P_{N_j}(z)$$

Finally, one can easily see that

$$\begin{aligned} E(N^m) &= \sum_{n=0}^{+\infty} n^m p_n = \sum_{n=0}^{+\infty} n^m (a_1 P(N_1 = n) + \cdots + a_k P(N_k = n)) \\ &= a_1 E(N_1^m) + \cdots + a_k E(N_k^m) \end{aligned}$$

Example 2.10. *Let N be a $2-$point mixture of the random variables N_1 and N_2 with weights p and $1-p$ then*

$$E(N) = pE(N_1) + (1-p)E(N_2)$$

and

$$\begin{aligned} Var(N) &= E(N^2) - (E(N))^2 \\ &= pE(N_1^2) + (1-p)E(N_2^2) - p^2(E(N_1))^2 - (1-p)^2(E(N_2))^2 + 2p(1-p)E(N_1)E(N_2) \\ &= pVar(N_1) + (1-p)Var(N_2) + p(1-p)(E(N_1) + E(N_2))^2 \end{aligned}$$

Example 2.11. *Let N be a $2-$ point mixture of a Bernoulli random variable with parameter q and geometric with parameter β where the weights are p and $1-p$. Then*

$$\begin{aligned}
P(N=0) &= p(1-q) + (1-p)\frac{1}{1+\beta} \\
P(N=1) &= pq + (1-p)\left(\frac{1}{1+\beta}\right)\left(\frac{\beta}{1-\beta}\right) \\
P(N=k) &= (1-p)\left(\frac{1}{1+\beta}\right)\left(\frac{\beta}{1-\beta}\right)^k, \quad \text{for } k = 2, 3, \cdots .
\end{aligned}$$

Example 2.12. *The zero-modified distribution can be created by using a two point mixture, of a degenerate distribution that places all probability at zero and a distribution with the original probability function. Indeed,*

$$\begin{aligned}
P_N^M(z) = \frac{p_0^M - p_0}{1-p_0} + \frac{1-p_0^M}{1-p_0}P_N(z) &= \frac{P_N(z) - p_0}{1-p_0} + p_0^M \frac{1-P_N^M}{1-p_0}P_N(z) \\
&= P_N^T(z) + p_0^M(-P_N^T + 1) = p_0^M \times 1 + (1-p_0^M)P_N^T(z)
\end{aligned}$$

2.5.2 Continuous mixture distributions

We have a continuous mixture distribution when the mixing distribution is continuous.

Definition 2.8. *The random variable N is a mixture of random variables when*

$$p_n = \int_{-\infty}^{+\infty} P(N = n|\Theta = \theta)u(\theta)d\theta.$$

The following proposition provides information on how to compute some quantities of N.

Proposition 2.7. *Let N be a mixture distribution. Then,*

$$\begin{aligned}
F_N(n) = E(F_{N|\Theta}(n)) &= \int_{-\infty}^{+\infty} P(N \leq n|\Theta = \theta)u(\theta)d\theta \\
E(N^k) &= E(E(N^k|\Theta)) \\
Var(N) &= E(Var(N|\Theta)) + Var(E(N|\Theta))
\end{aligned}$$

Proof.

$$\begin{aligned}
E(Var(N|\Theta)) + Var(E(N|\Theta)) &= E(E(N^2|\Theta) - (E(\mathbb{N}|\Theta))^2) + E((E(\mathbb{N}|\Theta))^2) - (E(E(\mathbb{N}|\Theta)))^2 \\
&= E(E(N^2|\Theta)) - (E(E(\mathbb{N}|\Theta)))^2 = E(N^2) - (E(N))^2 = Var(N)
\end{aligned}$$

∎

Example 2.13. *The probability generating function of a mixed Poisson distribution with a general mixing distribution Θ with distribution U is given by*

$$P(z) = E\left(e^{\Theta(z-1)}\right) = M_\Theta(z-1).$$

Frequently, the pgf of a mixed Poisson random variable is presented as

$$P(z) = E\left(e^{\lambda\Theta(z-1)}\right) = M_\Theta(\lambda(z-1)).$$

where, λ is a rescale parameter. Naturally, the two pgf are equivalent because $\lambda\Theta$ is itself a random variable, say $\tilde{\Theta}$.

Example 2.14. *Determine the p.f. of a mixed binomial with a beta mixing distribution (called binomial-beta).*

$$\begin{aligned}
p_k &= \int_0^1 \binom{m}{k} q^k(1-q)^{m-k}\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}q^{a-1}(1-q)^{b-1}dq = \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{\Gamma(m+1)}{\Gamma(k+1)\Gamma(m-k+1)}\int_0^1 q^{k+a-1}(1-q)^{m-k+b-1}dq = \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{\Gamma(m+1)}{\Gamma(k+1)\Gamma(m-k+1)}\frac{\Gamma(k+a)\Gamma(m-k+b)}{\Gamma(m+a+b)} \\
&\times \underbrace{\int_0^1 \frac{\Gamma(m+a+b)}{\Gamma(k+a)\Gamma(m-k+b)}q^{k+a-1}(1-q)^{m-k+b-1}dq =}_{=1} \\
&= \frac{\Gamma(a+b)\Gamma(m+1)}{\Gamma(m+a+b)}\frac{\Gamma(m-k+b)}{\Gamma(b)\Gamma(m-k+1)}\frac{\Gamma(k+a)}{\Gamma(a)\Gamma(k+1)} \\
&= \frac{\binom{a+k-1}{k}\binom{b+m-k-1}{m-k}}{\binom{a+b+m-1}{m}}, \ k = 0, 1, 2..
\end{aligned}$$

Example 2.15. *Show that the composition of a Poisson with the ETNB with $r = -0.5$ can be obtained as a mixture of the Poisson with the inverse Gaussian.*

Solution: *It is known that the $Poisson - ETNB$ with $r = -0.5$ has a pgf given by*

$$P(z) = e^{\lambda(P_2(z)-1)}$$

where $P_2(z)$ is the pgf of an extended truncated negative binomial with $r = -0.5$, i.e.

$$P_2(z) = \frac{(1-2\beta(z-1))^{1/2} - (1+2\beta)^{1/2}}{1-(1+2\beta)^{1/2}}$$

Therefore, straightforward computations result in

$$P(z) = e^{\lambda\left(\frac{(1+2\beta(1-z))^{1/2}-1}{1-(1+2\beta)^{1/2}}\right)}$$

As seen before, the pgf of a mixed Poisson with an inverse Gaussian mixing distribution is $M_3(z-1)$, where M_3 represents the mgf of an inverse Gaussian distribution. In general, the mgf of an inverse Gaussian with parameters μ, θ is

$$e^{\frac{\theta}{\mu}\left(1-\left(1-\frac{2\mu^2}{\theta}z\right)^{1/2}\right)} = e^{-\frac{\mu}{\beta}\left((1-2z\beta)^{1/2}-1\right)}, \quad \textit{for all } t < \frac{\theta}{2\mu^2}$$

if we set $\beta = \frac{\mu^2}{\theta}$. Consequently,

$$M_3(z-1) = e^{-\frac{\mu}{\beta}\left((1+2(1-z)\beta)^{1/2}-1\right)}$$

Fixing $\lambda = \frac{\mu}{\beta}((1+2\beta)^{1/2}-1)$ we get the result.

Before we finish this section, we will present a result establishing a relationship between a mixed Poisson random variable and a compound Poisson random variable.

Definition 2.9. *A distribution is said to be infinitely divisible if for $n \in \mathbb{N}$ its characteristic function can be written as*

$$\psi(z) = [\psi_n(z)]^n$$

where ψ_n is a characteristic function of some random variable. If the probability generating function exists, then we can replace the characteristic function by the probability generating function.

Theorem 2.9. *Suppose that $P(z)$ is the pgf of a mixed Poisson with an infinitely divisible mixing distribution. Then, there is a new pgf, $P_2(z)$ such that*

$$P(z) = e^{-\lambda(P_2(z)-1)}$$

that is a pgf of compounded Poisson distribution.

Examples of infinitely divisible distributions are the normal, gamma, Poisson, and negative binomial distributions. The binomial distribution is not infinitely divisible.

2.6 Mixed Poisson Processes

In this section, instead of assuming that the number of claims is a Poisson process with intensity λ, we suppose that λ is the result of the observation of a non-negative random variable, Λ. Let U be the cumulative distribution function of Λ, i.e.

$$U(\lambda) = Pr(\Lambda \leq \lambda).$$

The random variable Λ is called structure random variable and $U(\lambda)$ structure distribution.

Definition 2.10. *The unconditional counting process $\{N(t) : t \geq 0\}$, with $N(0) = 0$ and such that*

$$\begin{aligned} Pr(N(t+s) - N(s) = k) &= \int_0^\infty Pr(N(t+s) - N(s) = k|\Lambda = \lambda) dU(\lambda) \\ &= \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^k}{k!} dU(\lambda). \end{aligned}$$

is called a mixed Poisson process.

One can easily see that $Pr(N(t+s) - N(s) = k)$ only depends on the t, therefore, the increments are stationary. The random variable $N(t)$ has a mixed Poisson distribution

$$Pr(N(t) = k) = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^k}{k!} dU(\lambda).$$

On the other hand, one can notice that mixed Poisson process has not independent increments because

$$Pr(N(t_2) - N(t_1) = k_2, N(t_1) - N(t_0) = k_1) \neq Pr(N(t_2) - N(t_1) = k_2) P(N(t_1) - N(t_0) = k_1).$$

It is normal to consider that the mixed Poisson process is the Bayesian version of the Poisson process where U is the a priori distribution of the intensity of the process. The a posteriori distribution of the intensity is

$$U^*(x) = Pr(\Lambda \leq x | N(t) = k) = \frac{\int_0^x \lambda^k e^{-\lambda t} dU(\lambda)}{\int_0^\infty \lambda^k e^{-\lambda t} dU(\lambda)}$$

The probability generating function of the random variable $N(t)$ is

$$P_{N(t)}(z) = E\left(E\left(z^{N(t)}|\Lambda\right)\right) = E(e^{\Lambda t(z-1)}) = M_\Lambda(t(z-1)).$$

Similarly, the mgf and cgf are given by

$$M_{N(t)}(r) = M_\Lambda(t(e^r - 1)), \quad g_{N(t)}(s) = \ln(M_{N(t)}(s)) = \ln(M_\Lambda(t(e^s - 1))) = g_\Lambda(t(e^s - 1)).$$

In light of these results, one gets that

$$E(N(t)) = P'_{N(t)}(1) = tE(\Lambda), \quad E(N(t)(N(t) - 1)) = P''_{N(t)}(1) = t^2 E(\Lambda^2).$$

Therefore,

$$Var(N(t)) = E(N(t)(N(t) - 1)) + E(N) - (E(N(t)))^2 = t^2 Var(\Lambda) + tE(\Lambda).$$

The asymmetric coefficient is given by

$$\gamma_{N(t)} = \frac{tE[\Lambda] + 3t^2 Var[\Lambda] + t^3 \mu_3(\lambda)}{(Var(N(t)))^{3/2}},$$

where μ_3 is the third central moment.

2.6.1 The Polya process

The Polya process is a particular case of the mixed Poisson process when the structure variable follows a gamma distribution, i.e

$$u(\lambda) = \frac{1}{\Gamma(r)\beta^r} e^{-\lambda/\beta} \lambda^{r-1}, \quad \lambda > 0$$

where $r > 0$ is the shape parameter and β is the scale parameter. The mgf of Λ is

$$M_\Lambda(t) = (1 - \beta t)^{-r}, \quad t < 1/\beta.$$

The expected value, the variance and the third central moment is

$$E(\Lambda) = \beta r, \quad Var(\Lambda) = \beta^2 r \quad \text{and } E[(\Lambda - \mu)^3] = 2\beta^3 r$$

Taking into account that

$$P_{N(t)}(z) = M_\Lambda(t(z-1)) = (1 - \beta t(z-1))^{-r}$$

we can conclude that $N(t)$ follows a negative binomial with parameters r and βt. Finally,

$$E(N(t)) = \beta r t \quad \text{and} \quad Var(N(t)) = \beta^2 r t^2 + \beta r t$$

2.7 Effects of exposure on frequency

We should expect that the number of claims in a certain period of time, N, increases with the exposure (number of lives, number policies, square meters of insured buildings,...).

Suppose that the portfolio consists of n entities, each of them producing claims N_j in the period under consideration. Then $N = N_1 + N_2 + \ldots + N_n$. If we suppose that N_j are i.i.d., then

$$P_N(z) = [P_{N_1}(z)]^n.$$

If instead of n there were n^* entities, then N^*, would have probability generating function

$$P_{N^*}(z) = [P_{N_1}(z)]^{n^*} = [P_N(z)]^{n^*/n}.$$

If N is infinitely divisible, then N^* will have the same form as N but with different parameter.

Example 2.16. *Consider a health plan for a group of 300 teachers of a school, and suppose that the number of claims of the group is considered to follow a negative binomial with parameters* $r = 10$ *and* $\beta = 3$. *The distribution of the number of claims for another similar group of 450 teachers could still be considered negative binomial with the same* β *and* $r = 15 = 10 \times 450/300$. *To check this result, one may notice that*

$$P_{N^*} = \left[(1 - \beta(z-1))^{-r}\right]^{450/300} = (1 - \beta(z-1))^{-15r}.$$

2.8 Exercises

- Section 2.1: Exercises 6.2, and 6.3 from the Loss Models book (3rd edition), Exercise 1 from exam 03/06/2013, and Exercise 1 from exam 26/06/2013;
- Section 2.3: Exercise 6.32 from the Loss Models book and Exercises 1 from exam 25/06/2012;
- Section 2.4: Exercise 3 from exam 03/07/2017, Exercise 2 from exam 30/06/2014, Exercises 2 and 3 from exam, 03/06/2013, and 2 from exam 25/06/2012;
- Section 2.7: Exercise 3 from exam 02/06/2014

1. Let N be a counting distribution belonging to the $(a, b, 0)$ class of distributions such that
$$P(N = 0) = 0.1, \quad P(N = 1) = 0.3 \quad \text{and} \quad P(N = 2) = 0.3.$$
Find the parameters a and b as well as the distribution of N.

2. An insurance company has a portfolio with 30 life insurance policies, each with a 3.5% probability of loss in one year. The probabilities of loss are independent. Compute the probability that at least 4 or more policies generate losses in one year. On average, how often would 4 or more risks have losses in the same year?

3. Let N be a distribution from the $(a, b, 0)$ class of distributions such that
$$\frac{p_k}{p_{k-1}} = c + \frac{c}{k} \quad \text{and} \quad p_0 = \frac{1}{4}.$$
Find c and the distribution of N.

4. Assume that an insurance company has a portfolio with 20 independent risk. The distribution of the number of claims per year is a Poisson random variable. The average number of claims is 0.1, for 4 risks, 0.2 for 10 risks and 0.3 for the remaining risks. Compute the probability that the portfolio have at least 5 claims per year.

5. The number of claims of a given risk follows a Poisson process with intensity 2 per year.
   a) Compute the probability that in two years that risk generates less than 3 claims.
   b) If 4 claims were generated until the third year, what is the probability that 2 claims were generated until the second year.

6. For a certain risk, the total number of occurrences is modeled by a negative binomial with parameters r and β. However, the policy holder only reports occurrences whose losses are greater than a certain value, which occurs with probability p. Prove that the number of occurrences reported follows a negative binomial. Find the parameters.

7. Consider a Geometric distribution with $\beta = 3$. Compute p_k, p_k^T and p_k^M, for $k = 0, 1, 2, 3$, with $p_0 = 0.5$.

8. Prove that the probability generating function of a logarithmic distribution is given by

$$P_N(z) = 1 - \frac{\ln(1 + \beta(1 - z))}{\ln(1 + \beta)}.$$

9. Let N represent the logarithmic distribution. Prove that

$$E(N) = \frac{\beta}{\ln(1 + \beta)}, \quad Var(N) = \beta \frac{1 + \beta - \beta/\ln(1 + \beta)}{\ln(1 + \beta)}$$

10. Compute the expected value and variance of the zero-modified distributions.

11. Assume that the number of car accidents is modeled by the random variable N and M_i represents the number of claims if there are i accidents. The total number of claims is given by $S = M_0 + M_1 + \cdots M_N$, with $M = 0$ assuming that the random variables M_i are independent and identically distributed to M and

$$f_M(x) = \frac{1}{4}, \, x = 1, 2, 3, 4 \quad \text{and} \quad f_N(x) = \frac{1}{4}, \, x = 0, 1, 2, 3.$$

Compute $F_S(5)$.

12. Prove that the probability function of a Poisson random variable with a mixing continuous uniform random variable in the interval (a, b), with $a >$ and $b > 0$ is

$$p_k = \frac{P(a < X_k < b)}{b - a},$$

where X_k follows a Gamma distribution with parameters $\alpha = k + 1$ and $\theta = 1$.

13. Compute the expected value and variance of binomial distribution with parameters Q and m where Q follows a continuous uniform in the interval $(0, 1)$.

Transforms

Transform are useful instruments that characterize the distribution of random variables (when they exist). Often, the distribution of the sum of independent random variables can be obtained using the mgf or the pgf when we have a counting random variables. If $p_k = \Pr\{N = k\},\ k = 0, 1, 2, ...$ represents the probability function of a counting random variable N then, we can define the following generating functions:

a) Moment generating function

$$M_N(r) = E[e^{rN}] = \sum_{k=0}^{\infty} p_k e^{rk}$$

b) Probability generating function

$$P_N(z) = E[z^N] = \sum_{k=0}^{\infty} p_k z^k$$

c) Cumulant generating function

$$g_N(s) = \ln(M_N(r)) = \sum_{k=1}^{\infty} \kappa_k \frac{s^k}{k!}$$

The moment generating function, when exists, allows us to recover moments of order k, with $k = 1, 2, \cdots$, computing its derivative of order k at 0. Indeed, one can see that

$$M_N(r) = E[e^{rN}] = E\left[\sum_{k=0}^{\infty} \frac{(rN)^k}{k!}\right] = \sum_{k=0}^{\infty} E\left[N^k\right] \frac{r^k}{k!}.$$

Assuming that N is a counting variable, then the probability generating function allows us to recover all the probabilities. In fact, it is not difficult to check that

$$p_k = \frac{P_N^{(k)}(0)}{k!}.$$

Additionally, we can compute the kth factorial moment by using the probability generating function:

$$P_N^{(k)}(1^-) = E(N(N-1)\cdots(N-(k-1)))$$

.This means that the variance can be computed as $Var(N) = P_N''(1^-) - (P_N'(1^-))^2 + P_N'(1^-)$.

The cumulant generating function is such that

$$g_N(s) = E(N)s + \frac{Var(N)}{2}s^2 + \frac{E[(N-E(N))^3]}{6}s^3 + O(s^4).$$

To get this expansion, one can combine the fact that $\ln(1+z) = z - \frac{1}{2}z^2 + \frac{1}{3}z^3 + O(z^4)$ with the following expression

$$M_N(r) = \sum_{k=0}^{\infty} E\left[N^k\right]\frac{r^k}{k!} = 1 + E[N]r + E[N^2]\frac{r^2}{2!} + E[N^3]\frac{r^3}{3!} + O(r^4).$$

Gamma function

Gamma function is a generalization of the factorial and is defined by

$$\Gamma(\alpha) = \int_0^{\infty} e^{-x}x^{\alpha-1}dx, \quad (\alpha > 0).$$

There are some interesting properties about this function, namely the fact that

$$\begin{aligned}\Gamma(\alpha) &= \lambda^{\alpha}\int_0^{\infty} e^{-\lambda x}x^{\alpha-1}dx, \quad (\alpha > 0)\\ \Gamma(\alpha) &= \left[-e^{-x}x^{\alpha-1}\right]_0^{+\infty} + (\alpha-1)\int_0^{\infty} e^{-x}x^{\alpha-2}dx,\\ &= (\alpha-1)\Gamma(\alpha-1), \quad (\alpha > 0)\end{aligned}$$

It is a matter of computations to verify that

$$\Gamma(1) = \int_0^{\infty} e^{-x}dx = 1 \quad \text{and} \quad \Gamma(1/2) = \sqrt{\pi}.$$

As a consequence

$$\Gamma(n) = (n-1)(n-2)\cdots 2\Gamma(1) = (n-1)!.$$

Convolutions

In many situations in risk theory, we are interested in computing the probability/density function or the cumulative distribution function of a sum of independent random variables. The operation convolution provides an efficient way to compute that functions. Consider two independent random variable X and Y, the distribution function of their sum is

$$\begin{aligned}F_{X+Y}(z) = P(X+Y \leq z) = \int_{-\infty}^{\infty} P(X+Y \leq z|Y=y)dF_Y(y) &= \int_{-\infty}^{\infty} P(X \leq z-y)dF_Y(y)\\ = \int_{-\infty}^{\infty} F_X(z-y)dF_Y(y) = F_Y * F_X(z).\end{aligned}$$

Therefore, $F_Y * F_X(\cdot)$ represents the convolution F_Y with F_X. It is straightforward to see that the operation is commutative because $F_Y * F_X(z) = F_X * F_Y(z)$. The same operation can be applied to probability/density functions. If X and Y are both continuous, we get

$$F_{X+Y}(z) = \int_{-\infty}^{\infty} F_X(z-y) f_Y(y) dy \quad \text{and} \quad f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy.$$

On the other hand, when both are discrete then

$$F_{X+Y}(z) = \sum_{y \in D_Y} F_X(z-y) f_Y(y) \quad \text{and} \quad f_{X+Y}(z) = \sum_{y \in D_Y} f_X(z-y) f_Y(y) dy.$$

where D_Y is the set of discontinuity points of the cdf of Y. Convolutions can be used to compute the distribution of a sum of n independent random variables $X_1, X_2, \cdots, X_n$. In fact,

$$P(X_1 + \cdots + X_n \leq z) = (F_{X_1} * F_{X_2} * \cdots * F_{X_n})(z).$$

This implies that

$$\begin{aligned} P(X_1 + \cdots + X_n \leq z) &= \int_{-\infty}^{\infty} P(X_1 + \cdots + X_{n-1} \leq z - x_n) dF_{X_n}(x_n) \\ &= \int_{-\infty}^{\infty} (F_{X_1} * F_{X_2} * \cdots * F_{X_{n-1}})(z - x_n) dF_{X_n}(x_n), \end{aligned}$$

which provides a recursive formula to compute the cdf of a sum of random variables. Assuming that $X_1, X_2, \cdots, X_n$ are all independent and identically distributed to X, then

$$P(X_1 + \cdots + X_n \leq z) = F_X^{*n}(z),$$

where F_X^{*n} is the $n-$ fold convolution of F_X. Therefore, it is straightforward to see that

$$\begin{aligned} P(X_1 + \cdots + X_n \leq z) &= \int_{-\infty}^{\infty} P(X_1 + \cdots + X_{n-1} \leq z - x) dF_X(x) \\ &= \int_{-\infty}^{\infty} F_X^{*(n-1)}(z - x) dF_X(x), \end{aligned}$$

Example .17. Convolution of discrete random variables: *Assume that X_1, X_2 and X_3 are independent random variables such that*

$$f_{X_1}(x) = \begin{cases} 1/4, & x = 0 \\ 3/4, & x = 1 \end{cases}, \quad f_{X_2}(x) = \begin{cases} 1/2, & x = 0 \\ 1/2, & x = 2 \end{cases} \quad \text{and} \quad f_{X_3} = \begin{cases} 1/4, & x = 0 \\ 3/4, & x = 1 \end{cases}.$$

Compute the cumulative distribution function of $S = X_1 + X_2 + X_3$.

Solution:

z	$f_{X_1}(z)$	$*$	$f_{X_2}(z)$	$=$	$f_{X_1+X_2}(z)$	$*$	$f_{X_3}(z)$	$=$	$f_S(Z)$	$F_S(z)$
0	1/4		1/2		1/8		1/4		1/32	1/32
1	3/4		0		3/8		0		3/32	4/32
2	0		1/2		1/8		3/4		$3/32 + 1/32 = 4/32$	8/32
3	0		0		3/8		0		$9/32 + 3/32 = 12/32$	20/32
4	0		0		0		0		3/32	23/32
5	0		0		0		0		9/32	32/32

Example .18. Convolution of continuous random variables: *Assume that X is a uniform distribution in the interval $(0,1)$ and Y is a uniform distribution in the interval $(0,3)$. What is the CDF of $X+Y$?*

Solution: *The CDF of $X+Y$ is by definition*

$$P(X+Y \le z) = \int_{-\infty}^{\infty} F_X(z-y) f_Y(y) dy$$

where

$$F_X(x) = \begin{cases} 0, & x < 0 \\ x, & 0 < x \le 1 \\ 1, & x \ge 0 \end{cases} \quad \text{and} \quad f_Y(y) = \begin{cases} 1/3, & 0 < y < 3 \\ 0, & \text{otherwise} \end{cases}$$

If $0 < z < 1$ then

$$\int_{-\infty}^{\infty} F_X(z-y) f_Y(y) dy = \int_0^z (z-y)\frac{1}{3} dy = -\frac{1}{3}\left[\frac{(z-y)^2}{2}\right]_0^z = \frac{z^2}{6}$$

If $1 \le z < 3$ then

$$\begin{aligned}\int_{-\infty}^{\infty} F_X(z-y) f_Y(y) dy &= \int_0^{z-1} 1 \times \frac{1}{3} dy + \int_{z-1}^{z} (z-y) \times \frac{1}{3} dy \\ &= \frac{1}{3}(z-1) - \frac{1}{3}\left[\frac{(z-y)^2}{2}\right]_{z-1}^{z} = \frac{1}{3}\left(z - \frac{1}{2}\right)\end{aligned}$$

If $3 \le z < 4$ then

$$\begin{aligned}\int_{-\infty}^{\infty} F_X(z-y) f_Y(y) dy &= \int_0^{z-1} 1 \times \frac{1}{3} dy + \int_{z-1}^{3} (z-y) \times \frac{1}{3} dy \\ &= \frac{1}{3}(z-1) - \frac{1}{3}\left[\frac{(z-y)^2}{2}\right]_{z-1}^{3} \\ &= \frac{1}{3}\left(z - \frac{1}{2}\right) - \frac{1}{3} \times \frac{(z-3)^2}{2}.\end{aligned}$$

Therefore,

$$F_{X+Y}(z) = \begin{cases} 0, & z < 0 \\ \frac{z^2}{6}, & 0 \leq z < 1 \\ \frac{1}{3}\left(z - \frac{1}{2}\right), & 1 \leq z < 3 \\ \frac{1}{3}\left(\left(z - \frac{1}{2}\right) - \frac{(z-3)^2}{2}\right), & 3 \leq z < 4 \\ 1, & z \geq 4 \end{cases}$$

Example .19. *Let X_1, X_2 and X_3 be independent random variables with uniform discrete distribution taking values on the interval $\{1, 2, 3, 4\}$. Compute $P(X_1 + X_2 + X_3 \leq 7)$.*

Solution:

n	$f_X^{*1}(n)$	$f_X^{*2}(n)$	$f_X^{*3}(n)$	$F_{X_1+X_2+X_3}(n)$
1	1/4			
2	1/4	1/16		
3	1/4	2/16	1/64	1/64
4	1/4	3/16	3/64	4/64
5		4/16	6/64	10/64
6		3/16	10/64	20/64
7		2/16	12/64	32/64