Nonlife Actuarial Models

Chapter 1

Claim-Frequency Distribution

Learning Objectives

- Discrete distributions for modeling claim frequency
- Binomial, geometric, negative binomial and Poisson distributions
- The $(a, b, 0)$ and $(a, b, 1)$ classes of distributions
- Compound distribution
- Convolution
- Mixture distribution

1.2 Review of Statistics

- **Distribution function** (df) of random variable X

$$F_X(x) = \Pr(X \leq x). \tag{1.1}$$

- **Probability density function** (pdf) of continuous random variable

$$f_X(x) = \frac{dF_X(x)}{dx}. \tag{1.2}$$

- **Probability function** (pf) of discrete random variable

$$f_X(x) = \begin{cases} \Pr(X = x), & \text{if } x \in \Omega_X, \\ 0, & \text{otherwise.} \end{cases} \tag{1.3}$$

where Ω_X is the support of X

- **Moment generating function** (mgf), defined as

$$M_X(t) = \mathrm{E}(e^{tX}). \tag{1.6}$$

- Moments of X are obtainable from mgf by

$$M_X^r(t) = \frac{d^r M_X(t)}{dt^r} = \frac{d^r}{dt^r}\mathrm{E}(e^{tX}) = \mathrm{E}(X^r e^{tX}), \tag{1.7}$$

  so that

$$M_X^r(0) = \mathrm{E}(X^r) = \mu_r'. \tag{1.8}$$

- If $X_1, X_2, \cdots, X_n$ are **independently and identically distributed (iid)** random variables with mgf $M(t)$, and $X = X_1 + \cdots + X_n$, then the mgf of X is

$$M_X(t) = \mathrm{E}(e^{tX}) = \mathrm{E}\left(\prod_{i=1}^{n} e^{tX_i}\right) = \prod_{i=1}^{n} \mathrm{E}(e^{tX_i}) = [M(t)]^n . \qquad (1.9)$$

- **Probability generating function** (pgf), defined as $P_X(t) = \mathrm{E}(t^X)$,

$$P_X(t) = \sum_{x=0}^{\infty} t^x f_X(x), \qquad (1.13)$$

so that for X taking nonnegative integer values.

- We have

$$P_X^r(t) = \sum_{x=r}^{\infty} x(x-1)\cdots(x-r+1)t^{x-r} f_X(x) \qquad (1.14)$$

so that

$$f_X(r) = \frac{P_X^r(0)}{r!}$$

- Raw moments can be obtained by differentiating mgf,

- pf can be obtained by differentiating pgf.

- The mgf and pgf are related through the following equations

$$M_X(t) = P_X(e^t), \tag{1.11}$$

and

$$P_X(t) = M_X(\log t). \tag{1.12}$$

1.3 Some Discrete Distributions

(1) Binomial Distribution: $X \sim \mathcal{BN}(n, \theta)$ if

$$f_X(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, \qquad \text{for} \;\; x = 0, 1, \cdots, n, \tag{1.17}$$

- The mean and variance of X are

$$\mathrm{E}(X) = n\theta \qquad \text{and} \qquad \mathrm{Var}(X) = n\theta(1-\theta), \tag{1.19}$$

  so that the variance of X is always smaller than its mean.

- How do you prove these results?

- The mgf of X is

$$M_X(t) = (\theta e^t + 1 - \theta)^n, \tag{1.20}$$

and its pgf is

$$P_X(t) = (\theta t + 1 - \theta)^n. \tag{1.21}$$

- A recursive relationship for $f_X(x)$ is

$$f_X(x) = \left[\frac{(n - x + 1)\theta}{x(1 - \theta)}\right] f_X(x - 1) \tag{1.23}$$

(2) Geometric Distribution: $X \sim \mathcal{GM}(\theta)$ if

$$f_X(x) = \theta(1 - \theta)^x, \qquad \text{for} \;\; x = 0, 1, \cdots . \tag{1.24}$$

- The mean and variance of X are

$$\mathrm{E}(X) = \frac{1 - \theta}{\theta} \qquad \text{and} \qquad \mathrm{Var}(X) = \frac{1 - \theta}{\theta^2}, \tag{1.25}$$

- How do you prove these results?

- The mgf of X is

$$M_X(t) = \frac{\theta}{1 - (1-\theta)e^t}, \tag{1.26}$$

and its pgf is X

$$P_X(t) = \frac{\theta}{1 - (1-\theta)t}. \tag{1.27}$$

- The pf satisfies the following recursive relationship

$$f_X(x) = (1-\theta)\, f_X(x-1), \tag{1.28}$$

for $x = 1, 2, \cdots$, with starting value $f_X(0) = \theta$.

(3) Negative Binomial Distribution: $X \sim \mathcal{NB}(r, \theta)$ if

$$f_X(x) = \binom{x+r-1}{r-1}\theta^r(1-\theta)^x, \qquad \text{for} \quad x = 0, 1, \cdots, \tag{1.29}$$

- The mean and variance are

$$\mathrm{E}(X) = \frac{r(1-\theta)}{\theta} \qquad \text{and} \qquad \mathrm{Var}(X) = \frac{r(1-\theta)}{\theta^2}, \tag{1.30}$$

- The mgf of $\mathcal{NB}(r, \theta)$ is

$$M_X(t) = \left[\frac{\theta}{1-(1-\theta)e^t}\right]^r, \tag{1.31}$$

and its pgf is

$$P_X(t) = \left[\frac{\theta}{1-(1-\theta)t}\right]^r. \tag{1.32}$$

- May extend the parameter r to any positive number (not necessarily integer).

- The recursive formula of the pf is

$$f_X(x) = \left[\frac{(x+r-1)(1-\theta)}{x}\right] f_X(x-1), \tag{1.37}$$

with starting value

$$f_X(0) = \theta^r. \tag{1.38}$$

(4) Poisson Distribution: $X \sim \mathcal{PN}(\lambda)$, if the pf of X is given by

$$f_X(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad \text{for} \quad x = 0, 1, \cdots, \tag{1.39}$$

- The mean and variance of X are

$$\mathrm{E}(X) = \mathrm{Var}(X) = \lambda. \tag{1.40}$$

The mgf of X is

$$M_X(t) = \exp\left[\lambda(e^t - 1)\right], \tag{1.41}$$

and its pgf is

$$P_X(t) = \exp\left[\lambda(t - 1)\right]. \tag{1.42}$$

- Two important theorems of Poisson distribution

- **Theorem 1.1:** If $X_1, \cdots, X_n$ are independently distributed with $X_i \sim \mathcal{PN}(\lambda_i)$, for $i = 1, \cdots, n$, then $X = X_1 + \cdots + X_n$ is distributed as a Poisson with parameter $\lambda = \lambda_1 + \cdots + \lambda_n$.

- **Proof:** To prove this result, we make use of the mgf. Note that the mgf of X is

$$\begin{aligned}
M_X(t) &= \mathrm{E}(e^{tX}) \\
&= \mathrm{E}(e^{tX_1 + \cdots + tX_n}) \\
&= \mathrm{E}\left(\prod_{i=1}^{n} e^{tX_i}\right) \\
&= \prod_{i=1}^{n} \mathrm{E}(e^{tX_i}) \\
&= \prod_{i=1}^{n} \exp\left[\lambda_i(e^t - 1)\right] \\
&= \exp\left[(e^t - 1)\sum_{i=1}^{n} \lambda_i\right] \\
&= \exp\left[(e^t - 1)\lambda\right], \qquad (1.43)
\end{aligned}$$

which is the mgf of $\mathcal{PN}(\lambda)$.

- **Theorem 1.2:** Suppose an event A can be partitioned into m mutually exclusive and exhaustive events A_i, for $i = 1, \cdots, m$. Let X be the number of occurrences of A, and X_i be the number of occurrences of A_i, so that $X = X_1 + \cdots + X_m$. Let the probability of occurrence of A_i given A has occurred be p_i, i.e., $\Pr(A_i \mid A) = p_i$, with $\sum_{i=1}^{m} p_i = 1$. If $X \sim \mathcal{PN}(\lambda)$, then $X_i \sim \mathcal{PN}(\lambda_i)$, where $\lambda_i = \lambda p_i$. Furthermore, $X_1, \cdots, X_n$ are independently distributed.

- **Proof:** To prove this result, we first derive the marginal distribution of X_i. Given $X = x$, $X_i \sim \mathcal{BN}(x, p_i)$. Hence, the marginal pf of X_i is pf of $\mathcal{PN}(\lambda p_i)$. Then we show that the joint pf of $X_1, \cdots, X_m$ is the product of their marginal pf, so that $X_1, \cdots, X_m$ are independent.

Table A.1: Some discrete distributions

Distribution, parameters, notation and support	pf $f_X(x)$	mgf $M_X(t)$	Mean	Variance
Binomial $\mathcal{BN}(n,\theta)$ $x \in \{0,1,\cdots,n\}$	$\binom{n}{x}\theta^x(1-\theta)^{n-x}$	$(\theta e^t + 1 - \theta)^n$	$n\theta$	$n\theta(1-\theta)$
Poisson $\mathcal{PN}(\lambda)$ $x \in \{0,1,\cdots\}$	$\dfrac{\lambda^x e^{-\lambda}}{x!}$	$\exp\left[\lambda(e^t - 1)\right]$	λ	λ
Geometric $\mathcal{GM}(\theta)$ $x \in \{0,1,\cdots\}$	$\theta(1-\theta)^x$	$\dfrac{\theta}{1-(1-\theta)e^t}$	$\dfrac{1-\theta}{\theta}$	$\dfrac{1-\theta}{\theta^2}$
Negative binomial $\mathcal{NB}(r,\theta)$ $x \in \{0,1,\cdots\}$	$\binom{x+r-1}{r-1}\theta^r(1-\theta)^x$	$\left[\dfrac{\theta}{1-(1-\theta)e^t}\right]^r$	$\dfrac{r(1-\theta)}{\theta}$	$\dfrac{r(1-\theta)}{\theta^2}$

1.4 The $(a, b, 0)$ Class of Distributions

- **Definition 1.1:** A nonnegative discrete random variable X is in the $(a, b, 0)$ class if its pf $f_X(x)$ satisfies the following recursion

$$f_X(x) = \left(a + \frac{b}{x}\right) f_X(x-1), \qquad \text{for } x = 1, 2, \cdots, \tag{1.48}$$

where a and b are constants, with given $f_X(0)$.

- As an example, we consider the binomial distribution. Its pf can be written as follows

$$f_X(x) = \left[-\frac{\theta}{1-\theta} + \frac{\theta(n+1)}{(1-\theta)x}\right] f_X(x-1). \tag{1.49}$$

Thus, we let

$$a = -\frac{\theta}{1-\theta} \qquad \text{and} \qquad b = \frac{\theta(n+1)}{(1-\theta)}. \tag{1.50}$$

- Binomial, geometric, negative binomial and Poisson belong to the $(a, b, 0)$ class of distributions.

Table 1.2: The $(a, b, 0)$ class of distributions

Distribution	a	b	$f_X(0)$
Binomial: $\mathcal{BN}(n, \theta)$	$-\dfrac{\theta}{1-\theta}$	$\dfrac{\theta(n+1)}{1-\theta}$	$(1-\theta)^n$
Geometric: $\mathcal{GM}(\theta)$	$1-\theta$	0	θ
Negative binomial: $\mathcal{NB}(r, \theta)$	$1-\theta$	$(r-1)(1-\theta)$	θ^r
Poisson: $\mathcal{PN}(\lambda)$	0	λ	$e^{-\lambda}$

- It may be desirable to obtain a good fit of the distribution at zero claim based on empirical experience and yet preserve the shape to coincide with some simple parametric distributions.

- This can be achieved by specifying the zero probability while adopting the recursion to mimic a selected $(a, b, 0)$ distribution.

- Let $f_X(x)$ be the pf of a $(a, b, 0)$ distribution called the base distribution. We denote $f_X^M(x)$ as the pf that is a modification of $f_X(x)$.

- The probability at point zero, $f_X^M(0)$, is specified and $f_X^M(x)$ is related to $f_X(x)$ as follows

$$f_X^M(x) = cf_X(x), \qquad \text{for } x = 1, 2, \cdots, \tag{1.52}$$

where c is an appropriate constant.

- For $f_X^M(\cdot)$ to be a well defined pf, we must have

$$\begin{aligned} 1 &= f_X^M(0) + \sum_{x=1}^{\infty} f_X^M(x) \\ &= f_X^M(0) + c \sum_{x=1}^{\infty} f_X(x) \\ &= f_X^M(0) + c[1 - f_X(0)]. \end{aligned} \tag{1.53}$$

Thus, we conclude that

$$c = \frac{1 - f_X^M(0)}{1 - f_X(0)}. \tag{1.54}$$

Substituting c into equation (1.52) we obtain $f_X^M(x)$, for $x = 1, 2, \cdots$.

- Together with the given $f_X^M(0)$, we have a distribution with the desired zero-claim probability and the same recursion as the base $(a, b, 0)$ distribution.

- This is called the **zero-modified distribution** of the base $(a, b, 0)$ distribution.
- In particular, if $f_X^M(0) = 0$, the modified distribution cannot take value zero and is called the **zero-truncated distribution**.
- The zero-truncated distribution is a particular case of the zero-modified distribution.

1.5 Some Methods for Creating New Distributions

1.5.1 Compound distribution

- Let $X_1, \cdots, X_N$ be iid nonnegative integer-valued random variables, each distributed like X. We denote the sum of these random variables by S, so that

$$S = X_1 + \cdots + X_N. \tag{1.60}$$

- If N is itself a nonnegative integer-valued random variable distributed independently of $X_1, \cdots, X_N$, then S is said to have a **compound distribution**.

- The distribution of N is called the **primary distribution**, and the distribution of X is called the **secondary distribution**.

- We shall use the *primary-secondary* convention to name a compound distribution.

- Thus, if N is Poisson and X is geometric, S has a Poisson-geometric distribution.

- A **compound Poisson** distribution is a compound distribution where N is Poisson, for *any* secondary distribution.

- Consider the simple case where N has a degenerate distribution taking value n with probability 1. S is then the sum of n terms of X_i, where n is fixed. Suppose $n = 2$, so that $S = X_1 + X_2$.

- As the pf of X_1 and X_2 are $f_X(\cdot)$, the pf of S is given by

$$
\begin{aligned}
f_S(s) &= \Pr(X_1 + X_2 = s) \\
&= \sum_{x=0}^{s} \Pr(X_1 = x \text{ and } X_2 = s - x) \\
&= \sum_{x=0}^{s} f_X(s) f_X(s - x), \qquad (1.62)
\end{aligned}
$$

where the last line above is due to the independence of X_1 and X_2.

- The pf of S, $f_S(\cdot)$, is the **convolution** of $f_X(\cdot)$, denoted by $(f_X * f_X)(\cdot)$, i.e.,

$$
f_{X_1+X_2}(s) = (f_X * f_X)(s) = \sum_{x=0}^{s} f_X(x) f_X(s - x). \qquad (1.63)
$$

- Convolutions can be evaluated recursively. When $n = 3$, the 3-fold convolution is

$$\begin{aligned} f_{X_1+X_2+X_3}(s) &= (f_{X_1+X_2} * f_{X_3})(s) = \\ (f_{X_1} * f_{X_2} * f_{X_3})(s) &= (f_X * f_X * f_X)(s). \end{aligned} \tag{1.64}$$

- **Example 1.7:** Let the pf of X be $f_X(0) = 0.1$, $f_X(1) = 0$, $f_X(2) = 0.4$ and $f_X(3) = 0.5$. Find the 2-fold and 3-fold convolutions of X.

- **Solution:** We first compute the 2-fold convolution. For $s = 0$ and 1, the probabilities are

$$(f_X * f_X)(0) = f_X(0)f_X(0) = (0.1)(0.1) = 0.01,$$

and

$$(f_X * f_X)(1) = f_X(0)f_X(1) + f_X(1)f_X(0) = (0.1)(0) + (0)(0.1) = 0.$$

Other values are similarly computed as follows

$$(f_X * f_X)(2) = (0.1)(0.4) + (0.4)(0.1) = 0.08,$$

$$(f_X * f_X)(3) = (0.1)(0.5) + (0.5)(0.1) = 0.10,$$

$$(f_X * f_X)(4) = (0.4)(0.4) = 0.16,$$

$$(f_X * f_X)(5) = (0.4)(0.5) + (0.5)(0.4) = 0.40,$$

and

$$(f_X * f_X)(6) = (0.5)(0.5) = 0.25.$$

For the 3-fold convolution, we show some sample workings as follows

$$f_X^{*3}(0) = [f_X(0)] \left[f_X^{*2}(0) \right] = (0.1)(0.01) = 0.001,$$

$$f_X^{*3}(1) = [f_X(0)] \left[f_X^{*2}(1) \right] + [f_X(1)] \left[f_X^{*2}(0) \right] = 0,$$

and

$$
\begin{aligned}
f_X^{*3}(2) &= [f_X(0)]\left[f_X^{*2}(2)\right] + [f_X(1)]\left[f_X^{*2}(1)\right] + [f_X(2)]\left[f_X^{*2}(0)\right] \\
&= 0.012.
\end{aligned}
$$

- The results are summarized in Table 1.4
- We now return to the compound distribution in which the primary distribution N has a pf $f_N(\cdot)$. Using the total law of probability, we obtain the pf of the compound distribution S as

$$
\begin{aligned}
f_S(s) &= \sum_{n=0}^{\infty} \Pr(X_1 + \cdots + X_N = s \,|\, N = n)\, f_N(n) \\
&= \sum_{n=0}^{\infty} \Pr(X_1 + \cdots + X_n = s)\, f_N(n),
\end{aligned}
$$

in which the term $\Pr(X_1 + \cdots + X_n = s)$ can be calculated as the n-fold convolution of $f_X(\cdot)$.

- The evaluation of convolution is usually quite complex when n is large.

- **Theorem 1.4:** Let S be a compound distribution. If the primary distribution N has mgf $M_N(t)$ and the secondary distribution X has mgf $M_X(t)$, then the mgf of S is

$$M_S(t) = M_N[\log M_X(t)]. \tag{1.66}$$

If N has pgf $P_N(t)$ and X is nonnegative integer valued with pgf $P_X(t)$, then the pgf of S is

$$P_S(t) = P_N[P_X(t)]. \tag{1.67}$$

- **Proof:** The proof makes use of results in conditional expectation. We note that

$$
\begin{aligned}
M_S(t) &= \mathrm{E}\left(e^{tS}\right) \\
&= \mathrm{E}\left(e^{tX_1 + \cdots + tX_N}\right) \\
&= \mathrm{E}\left[\mathrm{E}\left(e^{tX_1 + \cdots + tX_N} \mid N\right)\right] \\
&= \mathrm{E}\left\{\left[\mathrm{E}\left(e^{tX}\right)\right]^N\right\} \\
&= \mathrm{E}\left\{[M_X(t)]^N\right\} \\
&= \mathrm{E}\left\{\left[e^{\log M_X(t)}\right]^N\right\} \\
&= M_N[\log M_X(t)]. \qquad (1.68)
\end{aligned}
$$

- Similarly we get $P_S(t) = P_N[P_X(t)]$.

- To compute the pf of S. We note that

$$
f_S(0) = P_S(0) = P_N[P_X(0)], \qquad (1.70)
$$

- Also, we have

$$f_S(1) = P'_S(0). \tag{1.71}$$

The derivative $P'_S(t)$ may be computed by differentiating $P_S(t)$ directly, or by the chain rule using the derivatives of $P_N(t)$ and $P_X(t)$, i.e.,

$$P'_S(t) = \{P'_N[P_X(t)]\}\, P'_X(t). \tag{1.72}$$

- **Example 1.8:** Let $N \sim \mathcal{PN}(\lambda)$ and $X \sim \mathcal{GM}(\theta)$. Calculate $f_S(0)$ and $f_S(1)$.

- **Solution:** The pgf of N is

$$P_N(t) = \exp[\lambda(t-1)],$$

and the pgf of X is

$$P_X(t) = \frac{\theta}{1-(1-\theta)t}.$$

The pgf of S is

$$P_S(t) = P_N[P_X(t)] = \exp\left[\lambda\left(\frac{\theta}{1-(1-\theta)t} - 1\right)\right],$$

from which we obtain

$$f_S(0) = P_S(0) = \exp\left[\lambda\left(\theta - 1\right)\right].$$

To calculate $f_S(1)$, we differentiate $P_S(t)$ directly to obtain

$$P_S'(t) = \exp\left[\lambda\left(\frac{\theta}{1-(1-\theta)t} - 1\right)\right] \frac{\lambda\theta(1-\theta)}{[1-(1-\theta)t]^2},$$

so that

$$f_S(1) = P_S'(0) = \exp\left[\lambda\left(\theta - 1\right)\right] \lambda\theta(1-\theta).$$

- The Panjer (1981) recursion is a recursive method for computing the pf of S, which applies to the case where the primary distribution N belongs to the $(a, b, 0)$ class.

- **Theorem 1.5:** If N belongs to the $(a, b, 0)$ class of distributions and X is a nonnegative integer-valued random variable, then the pf of S is given by the following recursion

$$f_S(s) = \frac{1}{1 - af_X(0)} \sum_{x=1}^{s} \left(a + \frac{bx}{s}\right) f_X(x) f_S(s-x), \qquad \text{for } s = 1, 2, \cdots, \tag{1.74}$$

with initial value $f_S(0)$ given by equation (1.70).

- **Proof:** See Dickson (2005), Section 4.5.2.

- The mean and variance of a compound distribution can be obtained from the means and variances of the primary and secondary distributions. Thus, the first two moments of the compound distribution can be obtained without computing its pf.

- **Theorem 1.6:** Consider the compound distribution. We denote $\mathrm{E}(N) = \mu_N$ and $\mathrm{Var}(N) = \sigma_N^2$, and likewise $\mathrm{E}(X) = \mu_X$ and $\mathrm{Var}(X) = \sigma_X^2$. The mean and variance of S are then given by

$$\mathrm{E}(S) = \mu_N \mu_X, \tag{1.75}$$

and

$$\mathrm{Var}(S) = \mu_N \sigma_X^2 + \sigma_N^2 \mu_X^2. \tag{1.76}$$

- **Proof:** We use the results in Appendix A.11 on conditional expectations to obtain

$$\mathrm{E}(S) = \mathrm{E}[\mathrm{E}(S \,|\, N)] = \mathrm{E}[\mathrm{E}(X_1 + \cdots + X_N \,|\, N)] = \mathrm{E}(N\mu_X) = \mu_N \mu_X. \tag{1.77}$$

From (A.115), we have

$$\begin{aligned}\mathrm{Var}(S) &= \mathrm{E}[\mathrm{Var}(S \,|\, N)] + \mathrm{Var}[\mathrm{E}(S \,|\, N)] \\ &= \mathrm{E}[N\sigma_X^2] + \mathrm{Var}(N\mu_X) \\ &= \mu_N\sigma_X^2 + \sigma_N^2\mu_X^2, \end{aligned} \tag{1.78}$$

which completes the proof.

- If S is a compound Poisson distribution with $N \sim \mathcal{PN}(\lambda)$, so that $\mu_N = \sigma_N^2 = \lambda$, then

$$\mathrm{Var}(S) = \lambda(\sigma_X^2 + \mu_X^2) = \lambda\,\mathrm{E}(X^2). \tag{1.79}$$

Proof of equation (1.78)

Given two random variables X and Y, the conditional variance $\mathrm{Var}(X \,|\, Y)$ is defined as $v(Y)$, where

$$v(y) = \mathrm{Var}(X \,|\, y) = \mathrm{E}\{[X - \mathrm{E}(X \,|\, y)]^2 \,|\, y\} = \mathrm{E}(X^2 \,|\, y) - [\mathrm{E}(X \,|\, y)]^2.$$

Thus, we have

$$\mathrm{Var}(X \mid Y) = \mathrm{E}(X^2 \mid Y) - [\mathrm{E}(X \mid Y)]^2,$$

which implies

$$\mathrm{E}(X^2 \mid Y) = \mathrm{Var}(X \mid Y) + [\mathrm{E}(X \mid Y)]^2.$$

Now we have

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 \\
&= \mathrm{E}[\mathrm{E}(X^2 \mid Y)] - [\mathrm{E}(X)]^2 \\
&= \mathrm{E}\{\mathrm{Var}(X \mid Y) + [\mathrm{E}(X \mid Y)]^2\} - [\mathrm{E}(X)]^2 \\
&= \mathrm{E}[\mathrm{Var}(X \mid Y)] + \mathrm{E}\{[\mathrm{E}(X \mid Y)]^2\} - [\mathrm{E}(X)]^2 \\
&= \mathrm{E}[\mathrm{Var}(X \mid Y)] + \mathrm{E}\{[\mathrm{E}(X \mid Y)]^2\} - \{\mathrm{E}[\mathrm{E}(X \mid Y)]\}^2 \\
&= \mathrm{E}[\mathrm{Var}(X \mid Y)] + \mathrm{Var}[\mathrm{E}(X \mid Y)].
\end{aligned}$$

- **Example 1.10:** Let $N \sim \mathcal{PN}(2)$ and $X \sim \mathcal{GM}(0.2)$. Calculate $\mathrm{E}(S)$ and $\mathrm{Var}(S)$. Repeat the calculation for $N \sim \mathcal{GM}(0.2)$ and $X \sim \mathcal{PN}(2)$.

- **Solution:** As $X \sim \mathcal{GM}(0.2)$, we have

$$\mu_X = \frac{1-\theta}{\theta} = \frac{0.8}{0.2} = 4,$$

and

$$\sigma_X^2 = \frac{1-\theta}{\theta^2} = \frac{0.8}{(0.2)^2} = 20.$$

If $N \sim \mathcal{PN}(2)$, we have $\mathrm{E}(S) = (4)(2) = 8$. Since N is Poisson, we have

$$\mathrm{Var}(S) = 2(20 + 4^2) = 72.$$

For $N \sim \mathcal{GM}(0.2)$ and $X \sim \mathcal{PN}(2)$, $\mu_N = 4$, $\sigma_N^2 = 20$, and $\mu_X =$

$\sigma_X^2 = 2$. Thus, $\mathrm{E}(S) = (4)(2) = 8$, and we have

$$\mathrm{Var}(S) = (4)(2) + (20)(4) = 88.$$

- We have seen that the sum of independently distributed Poisson distributions is also Poisson.

- It turns out that the sum of independently distributed *compound* Poisson distributions has also a *compound* Poisson distribution.

- **Theorem 1.7:** Suppose $S_1, \cdots, S_n$ have independently distributed compound Poisson distributions, where the Poisson parameter of S_i is λ_i and the pgf of the secondary distribution of S_i is $P_i(\cdot)$. Then $S = S_1 + \cdots + S_n$ has a compound Poisson distribution with Poisson parameter $\lambda = \lambda_1 + \cdots + \lambda_n$. The pgf of the secondary distribution of S is $P(t) = \sum_{i=1}^{n} w_i P_i(t)$, where $w_i = \lambda_i / \lambda$.

- Proof: The pgf of S is (see Example 1.11 for an application)

$$\begin{aligned} P_S(t) &= \mathrm{E}\left(t^{S_1 + \cdots + S_n}\right) \\ &= \prod_{i=1}^{n} P_{S_i}(t) \\ &= \prod_{i=1}^{n} \exp\left\{\lambda_i [P_i(t) - 1]\right\} \\ &= \exp\left\{\sum_{i=1}^{n} \lambda_i P_i(t) - \sum_{i=1}^{n} \lambda_i\right\} \\ &= \exp\left\{\sum_{i=1}^{n} \lambda_i P_i(t) - \lambda\right\} \\ &= \exp\left\{\lambda\left[\sum_{i=1}^{n} \frac{\lambda_i}{\lambda}[P_i(t)] - 1\right]\right\} \\ &= \exp\left\{\lambda[P(t) - 1]\right\}. \end{aligned} \tag{1.80}$$

1.5.2 Mixture distribution

- Let $X_1, \cdots, X_n$ be random variables with corresponding pf or pdf $f_{X_1}(\cdot), \cdots, f_{X_n}(\cdot)$ in the common support Ω. A new random variable X may be created with pf or pdf $f_X(\cdot)$ given by

$$f_X(x) = p_1 f_{X_1}(x) + \cdots + p_n f_{X_n}(x), \qquad x \in \Omega, \tag{1.82}$$

where $p_i \geq 0$ for $i = 1, \cdots, n$ and $\sum_{i=1}^n p_i = 1$.

- **Theorem 1.8:** The mean of X is

$$\mathrm{E}(X) = \mu = \sum_{i=1}^n p_i \mu_i, \tag{1.83}$$

and its variance is

$$\mathrm{Var}(X) = \sum_{i=1}^n p_i \left[(\mu_i - \mu)^2 + \sigma_i^2 \right]. \tag{1.84}$$

- **Example 1.12:** The claim frequency of a bad driver is distributed as $\mathcal{PN}(4)$, and the claim frequency of a good driver is distributed as $\mathcal{PN}(1)$. A town consists of 20% bad drivers and 80% good drivers. What is the mean and variance of the claim frequency of a randomly selected driver from the town?

- **Solution:** The mean of the claim frequency is

$$(0.2)(4) + (0.8)(1) = 1.6,$$

and its variance is

$$(0.2)\left[(4 - 1.6)^2 + 4\right] + (0.8)\left[(1 - 1.6)^2 + 1\right] = 3.04.$$

- The above can be generalized to continuous mixing.

1.5 Excel Computation Notes

Table 1.5: Some Excel functions

		Example	
X	Excel function	input	output
$\mathcal{BN}(n, \theta)$	`BINOMDIST(x1,x2,x3,ind)`	`BINOMDIST(4,10,0.3,FALSE)`	0.2001
	x1 $= x$	`BINOMDIST(4,10,0.3,TRUE)`	0.8497
	x2 $= n$		
	x3 $= \theta$		
$\mathcal{PN}(\lambda)$	`POISSON(x1,x2,ind)`	`POISSON(4,3.6,FALSE)`	0.1912
	x1 $= x$	`POISSON(4,3.6,TRUE)`	0.7064
	x2 $= \lambda$		
$\mathcal{NB}(r, \theta)$	`NEGBINOMDIST(x1,x2,x3)`	`NEGBINOMDIST(3,1,0.4)`	0.0864
	x1 $= x$	`NEGBINOMDIST(3,3,0.4)`	0.1382
	x2 $= r$		
	x3 $= \theta$		