Appendix A
An Inventory of Continuous Distributions

A.1 INTRODUCTION

Descriptions of the models are given starting in Section A.2. First, a few mathematical preliminaries are presented that indicate how the various quantities can be computed.

The incomplete gamma function[1] is given by

$$\Gamma(\alpha;x) = \frac{1}{\Gamma(\alpha)}\int_0^x t^{\alpha-1}e^{-t}\,dt, \quad \alpha > 0,\ x > 0,$$

$$\text{with } \Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}\,dt, \quad \alpha > 0.$$

A useful fact is $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$. Also, define

$$G(\alpha;x) = \int_x^\infty t^{\alpha-1}e^{-t}\,dt, \quad x > 0.$$

[1]Some references, such as [3], denote this integral $P(\alpha,x)$ and define $\Gamma(\alpha,x) = \int_x^\infty t^{\alpha-1}e^{-t}\,dt$. Note that this definition does not normalize by dividing by $\Gamma(\alpha)$. When using software to evaluate the incomplete gamma function, be sure to note how it is defined.

Loss Models: From Data to Decisions, 3rd. ed.
By Stuart A. Klugman, Harry H. Panjer, Gordon E. Willmot

At times we will need this integral for nonpositive values of α. Integration by parts produces the relationship

$$G(\alpha;x) = -\frac{x^\alpha e^{-x}}{\alpha} + \frac{1}{\alpha}G(\alpha+1;x).$$

This process can be repeated until the first argument of G is $\alpha+k$, a positive number. Then it can be evaluated from

$$G(\alpha+k;x) = \Gamma(\alpha+k)[1-\Gamma(\alpha+k;x)].$$

However, if α is a negative integer or zero, the value of $G(0;x)$ is needed. It is

$$G(0;x) = \int_x^\infty t^{-1}e^{-t}\,dt = E_1(x),$$

which is called the *exponential integral*. A series expansion for this integral is

$$E_1(x) = -0.57721566490153 - \ln x - \sum_{n=1}^{\infty}\frac{(-1)^n x^n}{n(n!)}.$$

When α is a positive integer, the incomplete gamma function can be evaluated exactly as given in the following theorem.

Theorem A.1 *For integer* α,

$$\Gamma(\alpha;x) = 1 - \sum_{j=0}^{\alpha-1}\frac{x^j e^{-x}}{j!}.$$

Proof: For $\alpha=1$, $\Gamma(1;x) = \int_0^x e^{-t}\,dt = 1-e^{-x}$, and so the theorem is true for this case. The proof is completed by induction. Assume it is true for $\alpha = 1,\ldots,n$. Then

$$\begin{aligned}
\Gamma(n+1;x) &= \frac{1}{n!}\int_0^x t^n e^{-t}\,dt \\
&= \frac{1}{n!}\left(-t^n e^{-t}\Big|_0^x + \int_0^x n t^{n-1}e^{-t}\,dt\right) \\
&= \frac{1}{n!}\left(-x^n e^{-x}\right) + \Gamma(n;x) \\
&= -\frac{x^n e^{-x}}{n!} + 1 - \sum_{j=0}^{n-1}\frac{x^j e^{-x}}{j!} \\
&= 1 - \sum_{j=0}^{n}\frac{x^j e^{-x}}{j!}.
\end{aligned}$$

□

The incomplete beta function is given by

$$\beta(a,b;x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^x t^{a-1}(1-t)^{b-1}\,dt, \quad a>0,\ b>0,\ 0<x<1,$$

where

$$\beta(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$$

is the beta function, and when $b<0$ (but $a > 1+\lfloor -b\rfloor$), repeated integration by parts produces

$$\begin{aligned}
\Gamma(a)\Gamma(b)\beta(a,b;x) &= -\Gamma(a+b)\left[\frac{x^{a-1}(1-x)^b}{b}\right. \\
&\quad + \frac{(a-1)x^{a-2}(1-x)^{b+1}}{b(b+1)} + \cdots \\
&\quad \left. + \frac{(a-1)\cdots(a-r)x^{a-r-1}(1-x)^{b+r}}{b(b+1)\cdots(b+r)}\right] \\
&\quad + \frac{(a-1)\cdots(a-r-1)}{b(b+1)\cdots(b+r)}\Gamma(a-r-1) \\
&\quad \times \Gamma(b+r+1)\beta(a-r-1,b+r+1;x),
\end{aligned}$$

where r is the smallest integer such that $b+r+1>0$. The first argument must be positive (that is, $a-r-1>0$).

Numerical approximations for both the incomplete gamma and the incomplete beta function are available in many statistical computing packages as well as in many spreadsheets because they are just the distribution functions of the gamma and beta distributions. The following approximations are taken from [3]. The suggestion regarding using different formulas for small and large x when evaluating the incomplete gamma function is from [144]. That reference also contains computer subroutines for evaluating these expressions. In particular, it provides an effective way of evaluating continued fractions.

For $x \le \alpha+1$ use the series expansion

$$\Gamma(\alpha;x) = \frac{x^\alpha e^{-x}}{\Gamma(\alpha)}\sum_{n=0}^{\infty}\frac{x^n}{\alpha(\alpha+1)\cdots(\alpha+n)}$$

while for $x > \alpha+1$, use the continued-fraction expansion

$$1-\Gamma(\alpha;x) = \frac{x^\alpha e^{-x}}{\Gamma(\alpha)}\cfrac{1}{x+\cfrac{1-\alpha}{1+\cfrac{1}{x+\cfrac{2-\alpha}{1+\cfrac{2}{x+\cdots}}}}}.$$

The incomplete gamma function can also be used to produce cumulative probabilities from the standard normal distribution. Let $\Phi(z) = \Pr(Z \le z)$, where Z has the standard normal distribution. Then, for $z \ge 0$, $\Phi(z) = 0.5 + \Gamma(0.5;z^2/2)/2$, while for $z<0$, $\Phi(z) = 1-\Phi(-z)$.

The incomplete beta function can be evaluated by the series expansion

$$\begin{aligned}
\beta(a,b;x) &= \frac{\Gamma(a+b)x^a(1-x)^b}{a\Gamma(a)\Gamma(b)} \\
&\quad \times \left[1 + \sum_{n=0}^{\infty}\frac{(a+b)(a+b+1)\cdots(a+b+n)}{(a+1)(a+2)\cdots(a+n+1)}x^{n+1}\right].
\end{aligned}$$

The gamma function itself can be found from

$$
\begin{aligned}
\ln\Gamma(\alpha) \doteq\ & (\alpha - \tfrac{1}{2})\ln\alpha - \alpha + \frac{\ln(2\pi)}{2} \\
& + \frac{1}{12\alpha} - \frac{1}{360\alpha^3} + \frac{1}{1{,}260\alpha^5} - \frac{1}{1{,}680\alpha^7} + \frac{1}{1{,}188\alpha^9} - \frac{691}{360{,}360\alpha^{11}} \\
& + \frac{1}{156\alpha^{13}} - \frac{3{,}617}{122{,}400\alpha^{15}} + \frac{43{,}867}{244{,}188\alpha^{17}} - \frac{174{,}611}{125{,}400\alpha^{19}}.
\end{aligned}
$$

For values of α above 10, the error is less than 10^{-19}. For values below 10, use the relationship

$$\ln\Gamma(\alpha) = \ln\Gamma(\alpha+1) - \ln\alpha.$$

The distributions are presented in the following way. First, the name is given along with the parameters. Many of the distributions have other names, which are noted in parentheses. Next the density function $f(x)$ and distribution function $F(x)$ are given. For some distributions, formulas for starting values are given. Within each family the distributions are presented in decreasing order with regard to the number of parameters. The Greek letters used are selected to be consistent. Any Greek letter that is not used in the distribution means that that distribution is a special case of one with more parameters but with the missing parameters set equal to 1. *Unless specifically indicated, all parameters must be positive.*

Except for two distributions, inflation can be recognized by simply inflating the scale parameter θ. That is, if X has a particular distribution, then cX has the same distribution type, with all parameters unchanged except θ is changed to $c\theta$. For the lognormal distribution, μ changes to $\mu + \ln(c)$ with σ unchanged, while for the inverse Gaussian, both μ and θ are multiplied by c.

For several of the distributions, starting values are suggested. They are not necessarily good estimators, just places from which to start an iterative procedure to maximize the likelihood or other objective function. These are found by either the methods of moments or percentile matching. The quantities used are

$$\text{Moments:}\quad m = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad t = \frac{1}{n}\sum_{i=1}^{n} x_i^2,$$

$$\text{Percentile matching:}\quad p = \text{25th percentile},\ q = \text{75th percentile}.$$

For grouped data or data that have been truncated or censored, these quantities may have to be approximated. Because the purpose is to obtain starting values and not a useful estimate, it is often sufficient to just ignore modifications. For three- and four-parameter distributions, starting values can be obtained by using estimates from a special case, then making the new parameters equal to 1. An all-purpose starting value rule (for when all else fails) is to set the scale parameter (θ) equal to the mean and set all other parameters equal to 2.

Risk measures may be calculated as follows. For $\mathrm{VaR}_p(X)$, the value-at-risk, solve the equation $p = F[\mathrm{Var}_p(X)]$. Where there are convenient explicit solutions, they are provided. For $\mathrm{TVaR}_p(X)$, the tail-value-at-risk, use the formula

$$\mathrm{TVaR}_p(X) = \mathrm{Var}_p(X) + \frac{\mathrm{E}(X) - \mathrm{E}[X \wedge \mathrm{Var}_p(X)]}{1-p}.$$

Explicit formulas are provided in a few cases.

All the distributions listed here (and many more) are discussed in great detail in [91]. In many cases, alternatives to maximum likelihood estimators are presented.

A.2 TRANSFORMED BETA FAMILY

A.2.1 Four-parameter distribution

A.2.1.1 *Transformed beta*—$\alpha, \theta, \gamma, \tau$ (generalized beta of the second kind, Pearson Type VI)[2]

$$
\begin{aligned}
f(x) &= \frac{\Gamma(\alpha+\tau)}{\Gamma(\alpha)\Gamma(\tau)} \frac{\gamma(x/\theta)^{\gamma\tau}}{x[1+(x/\theta)^\gamma]^{\alpha+\tau}}, \\
F(x) &= \beta(\tau, \alpha; u), \quad u = \frac{(x/\theta)^\gamma}{1+(x/\theta)^\gamma}, \\
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(\tau + k/\gamma)\Gamma(\alpha - k/\gamma)}{\Gamma(\alpha)\Gamma(\tau)}, \quad -\tau\gamma < k < \alpha\gamma, \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k \Gamma(\tau + k/\gamma)\Gamma(\alpha - k/\gamma)}{\Gamma(\alpha)\Gamma(\tau)} \beta(\tau + k/\gamma, \alpha - k/\gamma; u) \\
&\quad + x^k[1 - F(x)], \quad k > -\tau\gamma, \\
\text{Mode} &= \theta\left(\frac{\tau\gamma - 1}{\alpha\gamma + 1}\right)^{1/\gamma}, \quad \tau\gamma > 1, \text{ else } 0.
\end{aligned}
$$

A.2.2 Three-parameter distributions

A.2.2.1 *Generalized Pareto*—α, θ, τ (beta of the second kind)

$$
\begin{aligned}
f(x) &= \frac{\Gamma(\alpha+\tau)}{\Gamma(\alpha)\Gamma(\tau)} \frac{\theta^\alpha x^{\tau-1}}{(x+\theta)^{\alpha+\tau}}, \\
F(x) &= \beta(\tau, \alpha; u), \quad u = \frac{x}{x+\theta},
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(\tau + k)\Gamma(\alpha - k)}{\Gamma(\alpha)\Gamma(\tau)}, \quad -\tau < k < \alpha, \\
\mathrm{E}[X^k] &= \frac{\theta^k \tau(\tau+1)\cdots(\tau+k-1)}{(\alpha-1)\cdots(\alpha-k)} \quad \text{if } k \text{ is a positive integer}, \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k \Gamma(\tau + k)\Gamma(\alpha - k)}{\Gamma(\alpha)\Gamma(\tau)} \beta(\tau + k, \alpha - k; u), \\
&\quad + x^k[1 - F(x)], \quad k > -\tau, \\
\text{Mode} &= \theta\frac{\tau - 1}{\alpha + 1}, \quad \tau > 1, \text{ else } 0.
\end{aligned}
$$

[2] There is no inverse transformed beta distribution because the reciprocal has the same distribution, with α and τ interchanged and θ replaced with $1/\theta$.

A.2.2.2 Burr—α, θ, γ (Burr Type XII, Singh–Maddala)

$$
\begin{aligned}
f(x) &= \frac{\alpha\gamma(x/\theta)^\gamma}{x[1+(x/\theta)^\gamma]^{\alpha+1}}, \\
F(x) &= 1-u^\alpha, \quad u=\frac{1}{1+(x/\theta)^\gamma}, \\
\mathrm{VaR}_p(X) &= \theta[(1-p)^{-1/\alpha}-1]^{1/\gamma}, \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(1+k/\gamma)\Gamma(\alpha-k/\gamma)}{\Gamma(\alpha)}, \quad -\gamma<k<\alpha\gamma, \\
\mathrm{E}[(X\wedge x)^k] &= \frac{\theta^k\Gamma(1+k/\gamma)\Gamma(\alpha-k/\gamma)}{\Gamma(\alpha)}\beta(1+k/\gamma,\alpha-k/\gamma;1-u) \\
&\quad +x^k u^\alpha, \quad k>-\gamma, \\
\mathrm{Mode} &= \theta\left(\frac{\gamma-1}{\alpha\gamma+1}\right)^{1/\gamma}, \quad \gamma>1, \text{ else } 0.
\end{aligned}
$$

A.2.2.3 Inverse Burr—τ, θ, γ (Dagum)

$$
\begin{aligned}
f(x) &= \frac{\tau\gamma(x/\theta)^{\gamma\tau}}{x[1+(x/\theta)^\gamma]^{\tau+1}}, \\
F(x) &= u^\tau, \quad u=\frac{(x/\theta)^\gamma}{1+(x/\theta)^\gamma}, \\
\mathrm{VaR}_p(X) &= \theta(p^{-1/\tau}-1)^{-1/\gamma}, \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(\tau+k/\gamma)\Gamma(1-k/\gamma)}{\Gamma(\tau)}, \quad -\tau\gamma<k<\gamma, \\
\mathrm{E}[(X\wedge x)^k] &= \frac{\theta^k\Gamma(\tau+k/\gamma)\Gamma(1-k/\gamma)}{\Gamma(\tau)}\beta(\tau+k/\gamma,1-k/\gamma;u) \\
&\quad +x^k[1-u^\tau], \quad k>-\tau\gamma, \\
\mathrm{Mode} &= \theta\left(\frac{\tau\gamma-1}{\gamma+1}\right)^{1/\gamma}, \quad \tau\gamma>1, \text{ else } 0.
\end{aligned}
$$

A.2.3 Two-parameter distributions

A.2.3.1 Pareto—α, θ (Pareto Type II, Lomax)

$$
\begin{aligned}
f(x) &= \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}, \\
F(x) &= 1-\left(\frac{\theta}{x+\theta}\right)^\alpha, \\
\mathrm{VaR}_p(X) &= \theta[(1-p)^{-1/\alpha}-1], \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(k+1)\Gamma(\alpha-k)}{\Gamma(\alpha)}, \quad -1<k<\alpha, \\
\mathrm{E}[X^k] &= \frac{\theta^k k!}{(\alpha-1)\cdots(\alpha-k)} \quad \text{if } k \text{ is a positive integer}, \\
\mathrm{E}[X\wedge x] &= \frac{\theta}{\alpha-1}\left[1-\left(\frac{\theta}{x+\theta}\right)^{\alpha-1}\right], \quad \alpha\neq 1, \\
\mathrm{E}[X\wedge x] &= -\theta\ln\left(\frac{\theta}{x+\theta}\right), \quad \alpha=1, \\
\mathrm{TVaR}_p(X) &= \mathrm{VaR}_p(X)+\frac{\theta(1-p)^{-1/\alpha}}{\alpha-1}, \quad \alpha>1, \\
\mathrm{E}[(X\wedge x)^k] &= \frac{\theta^k\Gamma(k+1)\Gamma(\alpha-k)}{\Gamma(\alpha)}\beta[k+1,\alpha-k;x/(x+\theta)] \\
&\quad +x^k\left(\frac{\theta}{x+\theta}\right)^\alpha, \quad \text{all } k, \\
\mathrm{Mode} &= 0, \\
\hat{\alpha} &= 2\frac{t-m^2}{t-2m^2}, \quad \hat{\theta}=\frac{mt}{t-2m^2}.
\end{aligned}
$$

A.2.3.2 Inverse Pareto—τ, θ

$$
\begin{aligned}
f(x) &= \frac{\tau\theta x^{\tau-1}}{(x+\theta)^{\tau+1}}, \\
F(x) &= \left(\frac{x}{x+\theta}\right)^\tau, \\
\mathrm{VaR}_p(X) &= \theta[p^{-1/\tau}-1]^{-1}, \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(\tau+k)\Gamma(1-k)}{\Gamma(\tau)}, \quad -\tau<k<1, \\
\mathrm{E}[X^k] &= \frac{\theta^k(-k)!}{(\tau-1)\cdots(\tau+k)} \quad \text{if } k \text{ is a negative integer}, \\
\mathrm{E}[(X\wedge x)^k] &= \theta^k\tau\int_0^{x/(x+\theta)} y^{\tau+k-1}(1-y)^{-k}dy \\
&\quad +x^k\left[1-\left(\frac{x}{x+\theta}\right)^\tau\right], \quad k>-\tau, \\
\mathrm{Mode} &= \theta\frac{\tau-1}{2}, \quad \tau>1, \text{ else } 0.
\end{aligned}
$$

A.2.3.3 *Loglogistic*—γ, θ (Fisk)

$$\begin{aligned}
f(x) &= \frac{\gamma(x/\theta)^\gamma}{x[1+(x/\theta)^\gamma]^2}, \\
F(x) &= u, \quad u=\frac{(x/\theta)^\gamma}{1+(x/\theta)^\gamma}, \\
\mathrm{VaR}_p(X) &= \theta(p^{-1}-1)^{-1/\gamma}, \\
\mathrm{E}[X^k] &= \theta^k\Gamma(1+k/\gamma)\Gamma(1-k/\gamma), \quad -\gamma<k<\gamma, \\
\mathrm{E}[(X\wedge x)^k] &= \theta^k\Gamma(1+k/\gamma)\Gamma(1-k/\gamma)\beta(1+k/\gamma,1-k/\gamma;u) \\
&\quad +x^k(1-u), \quad k>-\gamma, \\
\mathrm{Mode} &= \theta\left(\frac{\gamma-1}{\gamma+1}\right)^{1/\gamma}, \quad \gamma>1, \text{ else } 0, \\
\hat{\gamma} &= \frac{2\ln(3)}{\ln(q)-\ln(p)}, \quad \hat{\theta}=\exp\left(\frac{\ln(q)+\ln(p)}{2}\right).
\end{aligned}$$

A.2.3.4 *Paralogistic*—α, θ This is a Burr distribution with $\gamma=\alpha$.

$$\begin{aligned}
f(x) &= \frac{\alpha^2(x/\theta)^\alpha}{x[1+(x/\theta)^\alpha]^{\alpha+1}}, \\
F(x) &= 1-u^\alpha, \quad u=\frac{1}{1+(x/\theta)^\alpha}, \\
\mathrm{VaR}_p(X) &= \theta[(1-p)^{-1/\alpha}-1]^{1/\alpha}, \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(1+k/\alpha)\Gamma(\alpha-k/\alpha)}{\Gamma(\alpha)}, \quad -\alpha<k<\alpha^2, \\
\mathrm{E}[(X\wedge x)^k] &= \frac{\theta^k\Gamma(1+k/\alpha)\Gamma(\alpha-k/\alpha)}{\Gamma(\alpha)}\beta(1+k/\alpha,\alpha-k/\alpha;1-u) \\
&\quad +x^k u^\alpha, \quad k>-\alpha, \\
\mathrm{Mode} &= \theta\left(\frac{\alpha-1}{\alpha^2+1}\right)^{1/\alpha}, \quad \alpha>1, \text{ else } 0.
\end{aligned}$$

Starting values can use estimates from the loglogistic (use γ for α) or Pareto (use α) distributions.

A.2.3.5 *Inverse paralogistic*—τ, θ This is an inverse Burr distribution with $\gamma=\tau$.

$$\begin{aligned}
f(x) &= \frac{\tau^2(x/\theta)^{\tau^2}}{x[1+(x/\theta)^\tau]^{\tau+1}}, \\
F(x) &= u^\tau, \quad u=\frac{(x/\theta)^\tau}{1+(x/\theta)^\tau},
\end{aligned}$$

$$\begin{aligned}
\mathrm{VaR}_p(X) &= \theta(p^{-1/\tau}-1)^{-1/\tau}, \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(\tau+k/\tau)\Gamma(1-k/\tau)}{\Gamma(\tau)}, \quad -\tau^2<k<\tau, \\
\mathrm{E}[(X\wedge x)^k] &= \frac{\theta^k\Gamma(\tau+k/\tau)\Gamma(1-k/\tau)}{\Gamma(\tau)}\beta(\tau+k/\tau,1-k/\tau;u) \\
&\quad +x^k[1-u^\tau], \quad k>-\tau^2, \\
\mathrm{Mode} &= \theta\,(\tau-1)^{1/\tau}, \quad \tau>1, \text{ else } 0.
\end{aligned}$$

Starting values can use estimates from the loglogistic (use γ for τ) or inverse Pareto (use τ) distributions.

A.3 TRANSFORMED GAMMA FAMILY

A.3.1 Three-parameter distributions

A.3.1.1 *Transformed gamma*—α, θ, τ (generalized gamma)

$$\begin{aligned}
f(x) &= \frac{\tau u^\alpha e^{-u}}{x\Gamma(\alpha)}, \quad u=(x/\theta)^\tau, \\
F(x) &= \Gamma(\alpha;u), \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(\alpha+k/\tau)}{\Gamma(\alpha)}, \quad k>-\alpha\tau, \\
\mathrm{E}[(X\wedge x)^k] &= \frac{\theta^k\Gamma(\alpha+k/\tau)}{\Gamma(\alpha)}\Gamma(\alpha+k/\tau;u) \\
&\quad +x^k[1-\Gamma(\alpha;u)], \quad k>-\alpha\tau, \\
\mathrm{Mode} &= \theta\left(\frac{\alpha\tau-1}{\tau}\right)^{1/\tau}, \quad \alpha\tau>1, \text{ else } 0.
\end{aligned}$$

A.3.1.2 Inverse transformed gamma—α, θ, τ (inverse generalized gamma)

$$\begin{aligned}
f(x) &= \frac{\tau u^\alpha e^{-u}}{x\Gamma(\alpha)}, \quad u = (\theta/x)^\tau, \\
F(x) &= 1 - \Gamma(\alpha; u), \\
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(\alpha - k/\tau)}{\Gamma(\alpha)}, \quad k < \alpha\tau, \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k \Gamma(\alpha - k/\tau)}{\Gamma(\alpha)}[1 - \Gamma(\alpha - k/\tau; u)] + x^k \Gamma(\alpha; u) \\
&= \frac{\theta^k G(\alpha - k/\tau; u)}{\Gamma(\alpha)} + x^k \Gamma(\alpha; u), \quad \text{all } k, \\
\text{Mode} &= \theta \left(\frac{\tau}{\alpha\tau + 1} \right)^{1/\tau}.
\end{aligned}$$

A.3.2 Two-parameter distributions

A.3.2.1 Gamma—α, θ (When $\alpha = n/2$ and $\theta = 2$, it is a chi-square distribution with n degrees of freedom.)

$$\begin{aligned}
f(x) &= \frac{(x/\theta)^\alpha e^{-x/\theta}}{x\Gamma(\alpha)}, \\
F(x) &= \Gamma(\alpha; x/\theta), \\
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(\alpha + k)}{\Gamma(\alpha)}, \quad k > -\alpha, \\
\mathrm{E}[X^k] &= \theta^k (\alpha + k - 1) \cdots \alpha \quad \text{if } k \text{ is a positive integer}, \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k \Gamma(\alpha + k)}{\Gamma(\alpha)} \Gamma(\alpha + k; x/\theta) + x^k [1 - \Gamma(\alpha; x/\theta)], \quad k > -\alpha, \\
\mathrm{E}[(X \wedge x)^k] &= \alpha(\alpha + 1) \cdots (\alpha + k - 1)\theta^k \Gamma(\alpha + k; x/\theta) \\
&\quad + x^k [1 - \Gamma(\alpha; x/\theta)] \quad \text{if } k \text{ is a positive integer}, \\
M(t) &= (1 - \theta t)^{-\alpha}, \quad t < 1/\theta, \\
\text{Mode} &= \theta(\alpha - 1), \quad \alpha > 1, \text{ else } 0, \\
\hat{\alpha} &= \frac{m^2}{t - m^2}, \quad \hat{\theta} = \frac{t - m^2}{m}.
\end{aligned}$$

A.3.2.2 Inverse gamma—α, θ (Vinci)

$$\begin{aligned}
f(x) &= \frac{(\theta/x)^\alpha e^{-\theta/x}}{x\Gamma(\alpha)}, \\
F(x) &= 1 - \Gamma(\alpha; \theta/x), \\
\mathrm{E}[X^k] &= \frac{\theta^k \Gamma(\alpha - k)}{\Gamma(\alpha)}, \quad k < \alpha, \\
\mathrm{E}[X^k] &= \frac{\theta^k}{(\alpha - 1) \cdots (\alpha - k)} \quad \text{if } k \text{ is a positive integer}, \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k \Gamma(\alpha - k)}{\Gamma(\alpha)}[1 - \Gamma(\alpha - k; \theta/x)] + x^k \Gamma(\alpha; \theta/x) \\
&= \frac{\theta^k G(\alpha - k; \theta/x)}{\Gamma(\alpha)} + x^k \Gamma(\alpha; \theta/x), \text{ all } k, \\
\text{Mode} &= \theta/(\alpha + 1), \\
\hat{\alpha} &= \frac{2t - m^2}{t - m^2}, \quad \hat{\theta} = \frac{mt}{t - m^2}.
\end{aligned}$$

A.3.2.3 Weibull—θ, τ

$$\begin{aligned}
f(x) &= \frac{\tau (x/\theta)^\tau e^{-(x/\theta)^\tau}}{x}, \\
F(x) &= 1 - e^{-(x/\theta)^\tau}, \\
\mathrm{VaR}_p(X) &= \theta[-\ln(1 - p)]^{1/\tau}, \\
\mathrm{E}[X^k] &= \theta^k \Gamma(1 + k/\tau), \quad k > -\tau, \\
\mathrm{E}[(X \wedge x)^k] &= \theta^k \Gamma(1 + k/\tau) \Gamma[1 + k/\tau; (x/\theta)^\tau] + x^k e^{-(x/\theta)^\tau}, \quad k > -\tau, \\
\text{Mode} &= \theta \left(\frac{\tau - 1}{\tau} \right)^{1/\tau}, \quad \tau > 1, \text{ else } 0, \\
\hat{\theta} &= \exp\left(\frac{g \ln(p) - \ln(q)}{g - 1} \right), \quad g = \frac{\ln(\ln(4))}{\ln(\ln(4/3))}, \\
\hat{\tau} &= \frac{\ln(\ln(4))}{\ln(q) - \ln(\hat{\theta})}.
\end{aligned}$$

A.3.2.4 *Inverse Weibull—θ, τ* (log-Gompertz)

$$
\begin{aligned}
f(x) &= \frac{\tau(\theta/x)^\tau e^{-(\theta/x)^\tau}}{x}, \\
F(x) &= e^{-(\theta/x)^\tau}, \\
\mathrm{VaR}_p(X) &= \theta(-\ln p)^{-1/\tau}, \\
\mathrm{E}[X^k] &= \theta^k\Gamma(1-k/\tau), \quad k<\tau, \\
\mathrm{E}[(X\wedge x)^k] &= \theta^k\Gamma(1-k/\tau)\{1-\Gamma[1-k/\tau;(\theta/x)^\tau]\} \\
&\quad + x^k\left[1-e^{-(\theta/x)^\tau}\right], \\
&= \theta^k G[1-k/\tau;(\theta/x)^\tau] + x^k\left[1-e^{-(\theta/x)^\tau}\right], \quad \text{all } k, \\
\text{Mode} &= \theta\left(\frac{\tau}{\tau+1}\right)^{1/\tau}, \\
\hat{\theta} &= \exp\left(\frac{g\ln(q)-\ln(p)}{g-1}\right), \quad g=\frac{\ln(\ln(4))}{\ln(\ln(4/3))}, \\
\hat{\tau} &= \frac{\ln(\ln(4))}{\ln(\hat{\theta})-\ln(p)}.
\end{aligned}
$$

A.3.3 One-parameter distributions

A.3.3.1 *Exponential—θ*

$$
\begin{aligned}
f(x) &= \frac{e^{-x/\theta}}{\theta}, \\
F(x) &= 1-e^{-x/\theta}, \\
\mathrm{VaR}_p(X) &= -\theta\ln(1-p), \\
\mathrm{E}[X^k] &= \theta^k\Gamma(k+1), \quad k>-1,
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}[X^k] &= \theta^k k! \quad \text{if } k \text{ is a positive integer}, \\
\mathrm{E}[X\wedge x] &= \theta(1-e^{-x/\theta}), \\
\mathrm{TVaR}_p(X) &= -\theta\ln(1-p)+\theta, \\
\mathrm{E}[(X\wedge x)^k] &= \theta^k\Gamma(k+1)\Gamma(k+1;x/\theta)+x^k e^{-x/\theta}, \quad k>-1, \\
\mathrm{E}[(X\wedge x)^k] &= \theta^k k!\Gamma(k+1;x/\theta)+x^k e^{-x/\theta} \quad \text{if } k>-1 \text{ is an integer}, \\
M(t) &= (1-\theta t)^{-1}, \quad t<1/\theta, \\
\text{Mode} &= 0, \\
\hat{\theta} &= m.
\end{aligned}
$$

A.3.3.2 *Inverse exponential—θ*

$$
\begin{aligned}
f(x) &= \frac{\theta e^{-\theta/x}}{x^2}, \\
F(x) &= e^{-\theta/x}, \\
\mathrm{VaR}_p(X) &= \theta(-\ln p)^{-1}, \\
\mathrm{E}[X^k] &= \theta^k\Gamma(1-k), \quad k<1, \\
\mathrm{E}[(X\wedge x)^k] &= \theta^k G(1-k;\theta/x)+x^k(1-e^{-\theta/x}), \quad \text{all } k, \\
\text{Mode} &= \theta/2, \\
\hat{\theta} &= -q\ln(3/4).
\end{aligned}
$$

A.4 DISTRIBUTIONS FOR LARGE LOSSES

The general form of most of these distribution has probability starting or ending at an arbitrary location. The versions presented here all use zero for that point. The distribution can always be shifted to start or end elsewhere.

A.4.1 Extreme value distributions

A.4.1.1 *Gumbel—θ, μ* (μ can be negative)

$$
\begin{aligned}
f(x) &= \frac{1}{\theta}\exp(-y)\exp\left[-\exp(-y)\right], \; y=\frac{x-\mu}{\theta}, \quad -\infty<x<\infty, \\
F(x) &= \exp\left[-\exp(-y)\right], \\
\mathrm{VaR}_p(X) &= \mu+\theta[-\ln(-\ln p)], \\
M(t) &= e^{\mu t}\Gamma(1-\theta t), \; t<1/\theta, \\
\mathrm{E}[X] &= \mu+0.577215664901538\theta, \\
\mathrm{Var}(X) &= \frac{\pi^2\theta^2}{6}.
\end{aligned}
$$

A.4.1.2 *Frechet—α, θ* This is the inverse Weibull distribution of Section A.3.2.4.

$$
\begin{aligned}
f(x) &= \frac{\alpha(x/\theta)^{-\alpha}e^{-(x/\theta)^{-\alpha}}}{x}, \\
F(x) &= e^{-(x/\theta)^{-\alpha}},
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{VaR}_p(X) &= \theta(-\ln p)^{1/\alpha}, \\
\mathrm{E}[X^k] &= \theta^k\Gamma(1-k/\alpha), \quad k<\alpha, \\
\mathrm{E}[(X\wedge x)^k] &= \theta^k\Gamma(1-k/\alpha)\{1-\Gamma[1-k/\alpha;(x/\theta)^{-\alpha}]\} \\
&\quad + x^k\left[1-e^{-(x/\theta)^{-\alpha}}\right], \\
&= \theta^k G[1-k/\alpha;(x/\theta)^{-\alpha}]+x^k\left[1-e^{-(x/\theta)^{-\alpha}}\right], \quad \text{all } k.
\end{aligned}
$$

A.4.1.3 Weibull—α, θ[3]

$$
\begin{aligned}
f(x) &= -\frac{\alpha(-x/\theta)^\alpha e^{-(-x/\theta)^\alpha}}{x}, \quad x \le 0, \\
F(x) &= e^{-(-x/\theta)^\alpha}, \quad x \le 0, \\
\mathrm{E}[X^k] &= (-1)^k \theta^k \Gamma(1+k/\alpha), \quad k > -\alpha, \ k \text{ an integer}, \\
\text{Mode} &= -\theta\left(\frac{\alpha-1}{\alpha}\right)^{1/\alpha}, \quad \alpha > 1, \text{ else } 0.
\end{aligned}
$$

A.4.2 Generalized Pareto distributions

A.4.2.1 *Generalized Pareto*—γ, θ This is the Pareto distribution of Section A.2.3.1 with α replaced by $1/\gamma$ and θ replaced by $\alpha\theta$.

$$F(x) = 1 - \left(1 + \gamma\frac{x}{\theta}\right)^{-1/\gamma}, \ x \ge 0.$$

A.4.2.2 *Exponential*—θ This is the same as the exponential distribution of Section A.3.3.1 and is the limiting case of the above distribution as $\gamma \to 0$.

A.4.2.3 *Pareto*—γ, θ This is the single-parameter Pareto distribution of Section A.5.1.4. From the above distribution, shift the probability to start at θ.

A.4.2.4 *Beta*—α, θ This is the beta distribution of Section A.6.1.2 with $a = 1$.

A.5 OTHER DISTRIBUTIONS

A.5.1.1 Lognormal—μ, σ (μ can be negative)

$$
\begin{aligned}
f(x) &= \frac{1}{x\sigma\sqrt{2\pi}}\exp(-z^2/2) = \phi(z)/(\sigma x), \quad z = \frac{\ln x - \mu}{\sigma}, \\
F(x) &= \Phi(z),
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}[X^k] &= \exp\left(k\mu + \tfrac{1}{2}k^2\sigma^2\right), \\
\mathrm{E}[(X \wedge x)^k] &= \exp\left(k\mu + \tfrac{1}{2}k^2\sigma^2\right)\Phi\left(\frac{\ln x - \mu - k\sigma^2}{\sigma}\right) + x^k[1 - F(x)], \\
\text{Mode} &= \exp(\mu - \sigma^2), \\
\hat{\sigma} &= \sqrt{\ln(t) - 2\ln(m)}, \quad \hat{\mu} = \ln(m) - \tfrac{1}{2}\hat{\sigma}^2.
\end{aligned}
$$

[3] This is not the same Weibull distribution as in Section A.3.2.3. It is the negative of a Weibull distribution.

A.5.1.2 Inverse Gaussian—μ, θ

$$
\begin{aligned}
f(x) &= \left(\frac{\theta}{2\pi x^3}\right)^{1/2}\exp\left(-\frac{\theta z^2}{2x}\right), \quad z = \frac{x-\mu}{\mu}, \\
F(x) &= \Phi\left[z\left(\frac{\theta}{x}\right)^{1/2}\right] + \exp\left(\frac{2\theta}{\mu}\right)\Phi\left[-y\left(\frac{\theta}{x}\right)^{1/2}\right], \quad y = \frac{x+\mu}{\mu}, \\
\mathrm{E}[X] &= \mu, \quad Var[X] = \mu^3/\theta, \\
\mathrm{E}[X^k] &= \sum_{n=0}^{k-1}\frac{(k+n-1)!}{(k-n-1)!n!}\frac{\mu^{n+k}}{(2\theta)^n}, \quad k = 1, 2, \ldots, \\
\mathrm{E}[X \wedge x] &= x - \mu z\Phi\left[z\left(\frac{\theta}{x}\right)^{1/2}\right] - \mu y\exp(2\theta/\mu)\Phi\left[-y\left(\frac{\theta}{x}\right)^{1/2}\right], \\
M(t) &= \exp\left[\frac{\theta}{\mu}\left(1 - \sqrt{1 - \frac{2\mu^2}{\theta}t}\right)\right], \quad t < \frac{\theta}{2\mu^2}, \\
\hat{\mu} &= m, \quad \hat{\theta} = \frac{m^3}{t - m^2}.
\end{aligned}
$$

A.5.1.3 *log-t*—r, μ, σ (μ can be negative) Let Y have a t distribution with r degrees of freedom. Then $X = \exp(\sigma Y + \mu)$ has the log-t distribution. Positive moments do not exist for this distribution. Just as the t distribution has a heavier tail than the normal distribution, this distribution has a heavier tail than the lognormal distribution.

$$
\begin{aligned}
f(x) &= \frac{\Gamma\left(\frac{r+1}{2}\right)}{x\sigma\sqrt{\pi r}\,\Gamma\left(\frac{r}{2}\right)\left[1 + \frac{1}{r}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right]^{(r+1)/2}}, \\
F(x) &= F_r\left(\frac{\ln x - \mu}{\sigma}\right) \text{ with } F_r(t) \text{ the cdf of a } t \text{ distribution with } r \text{ df}, \\
F(x) &= \begin{cases} \frac{1}{2}\beta\left[\frac{r}{2}, \frac{1}{2}; \frac{r}{r + \left(\frac{\ln x - \mu}{\sigma}\right)^2}\right], & 0 < x \le e^\mu, \\ 1 - \frac{1}{2}\beta\left[\frac{r}{2}, \frac{1}{2}; \frac{r}{r + \left(\frac{\ln x - \mu}{\sigma}\right)^2}\right], & x \ge e^\mu. \end{cases}
\end{aligned}
$$

A.5.1.4 *Single-parameter Pareto*—α, θ

$$
\begin{aligned}
f(x) &= \frac{\alpha\theta^\alpha}{x^{\alpha+1}}, \quad x > \theta, \\
F(x) &= 1 - \left(\frac{\theta}{x}\right)^\alpha, \quad x > \theta, \\
\mathrm{VaR}_p(X) &= \theta(1-p)^{-1/\alpha}, \\
\mathrm{E}[X^k] &= \frac{\alpha\theta^k}{\alpha-k}, \quad k < \alpha, \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\alpha\theta^k}{\alpha-k} - \frac{k\theta^\alpha}{(\alpha-k)x^{\alpha-k}}, \; x \ge \theta, \\
\mathrm{TVaR}_p(X) &= \frac{\alpha\theta(1-p)^{-1/\alpha}}{\alpha-1}, \quad \alpha > 1, \\
\text{Mode} &= \theta, \\
\hat{\alpha} &= \frac{m}{m-\theta}.
\end{aligned}
$$

Note: Although there appear to be two parameters, only α is a true parameter. The value of θ must be set in advance.

A.6 DISTRIBUTIONS WITH FINITE SUPPORT

For these two distributions, the scale parameter θ is assumed known.

A.6.1.1 *Generalized beta*—a, b, θ, τ

$$
\begin{aligned}
f(x) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^a(1-u)^{b-1}\frac{\tau}{x}, \quad 0 < x < \theta, \quad u = (x/\theta)^\tau, \\
F(x) &= \beta(a, b; u), \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(a+b)\Gamma(a+k/\tau)}{\Gamma(a)\Gamma(a+b+k/\tau)}, \quad k > -a\tau, \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k\Gamma(a+b)\Gamma(a+k/\tau)}{\Gamma(a)\Gamma(a+b+k/\tau)}\beta(a+k/\tau, b; u) + x^k[1-\beta(a, b; u)].
\end{aligned}
$$

A.6.1.2 *beta*—a, b, θ The case $\theta = 1$ has no special name, but is the commonly used version of this distribution.

$$
\begin{aligned}
f(x) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^a(1-u)^{b-1}\frac{1}{x}, \quad 0 < x < \theta, \quad u = x/\theta, \\
F(x) &= \beta(a, b; u), \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(a+b)\Gamma(a+k)}{\Gamma(a)\Gamma(a+b+k)}, \quad k > -a, \\
\mathrm{E}[X^k] &= \frac{\theta^k a(a+1)\cdots(a+k-1)}{(a+b)(a+b+1)\cdots(a+b+k-1)} \quad \text{if } k \text{ is a positive integer}, \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k a(a+1)\cdots(a+k-1)}{(a+b)(a+b+1)\cdots(a+b+k-1)}\beta(a+k, b; u) \\
&\quad + x^k[1-\beta(a, b; u)], \\
\hat{a} &= \frac{\theta m^2 - mt}{\theta t - \theta m^2}, \quad \hat{b} = \frac{(\theta m - t)(\theta - m)}{\theta t - \theta m^2}.
\end{aligned}
$$

Appendix B
An Inventory of Discrete Distributions

B.1 INTRODUCTION

The 16 models presented in this appendix fall into three classes. The divisions are based on the algorithm by which the probabilities are computed. For some of the more familiar distributions these formulas will look different from the ones you may have learned, but they produce the same probabilities. After each name, the parameters are given. All parameters are positive unless otherwise indicated. In all cases, p_k is the probability of observing k losses.

For finding moments, the most convenient form is to give the factorial moments. The jth factorial moment is $\mu_{(j)} = \mathrm{E}[N(N-1)\cdots(N-j+1)]$. We have $\mathrm{E}[N] = \mu_{(1)}$ and $\mathrm{Var}(N) = \mu_{(2)} + \mu_{(1)} - \mu_{(1)}^2$.

The estimators presented are not intended to be useful estimators but, rather, provide starting values for maximizing the likelihood (or other) function. For determining starting values, the following quantities are used (where n_k is the observed frequency at k [if, for the last entry, n_k represents the number of observations at k or more, assume it was at exactly k] and n is the sample size):

Loss Models: From Data to Decisions, 3rd. ed.
By Stuart A. Klugman, Harry H. Panjer, Gordon E. Willmot

$$\hat{\mu} = \frac{1}{n}\sum_{k=1}^{\infty} kn_k, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{k=1}^{\infty} k^2 n_k - \hat{\mu}^2.$$

When the method of moments is used to determine the starting value, a circumflex (e.g., $\hat{\lambda}$) is used. For any other method, a tilde (e.g., $\tilde{\lambda}$) is used. When the starting value formulas do not provide admissible parameter values, a truly crude guess is to set the product of all λ and β parameters equal to the sample mean and set all other parameters equal to 1. If there are two λ or β parameters, an easy choice is to set each to the square root of the sample mean.

The last item presented is the probability generating function,

$$P(z) = \mathrm{E}[z^N].$$

B.2 THE $(a, b, 0)$ CLASS

The distributions in this class have support on $0, 1, \ldots$. For this class, a particular distribution is specified by setting p_0 and then using $p_k = (a + b/k)p_{k-1}$. Specific members are created by setting p_0, a, and b. For any member, $\mu_{(1)} = (a+b)/(1-a)$, and for higher j, $\mu_{(j)} = (aj+b)\mu_{(j-1)}/(1-a)$. The variance is $(a+b)/(1-a)^2$.

B.2.1.1 Poisson—λ

$$\begin{aligned}
p_0 &= e^{-\lambda}, \quad a = 0, \quad b = \lambda, \\
p_k &= \frac{e^{-\lambda}\lambda^k}{k!}, \\
\mathrm{E}[N] &= \lambda, \quad \mathrm{Var}[N] = \lambda, \\
\hat{\lambda} &= \hat{\mu}, \\
P(z) &= e^{\lambda(z-1)}.
\end{aligned}$$

B.2.1.2 Geometric—β

$$\begin{aligned}
p_0 &= \frac{1}{1+\beta}, \quad a = \frac{\beta}{1+\beta}, \quad b = 0, \\
p_k &= \frac{\beta^k}{(1+\beta)^{k+1}}, \\
\mathrm{E}[N] &= \beta, \quad \mathrm{Var}[N] = \beta(1+\beta), \\
\hat{\beta} &= \hat{\mu}, \\
P(z) &= [1 - \beta(z-1)]^{-1}.
\end{aligned}$$

This is a special case of the negative binomial with $r = 1$.

B.2.1.3 Binomial—q, m, $(0 < q < 1, m \text{ an integer})$

$$\begin{aligned}
p_0 &= (1-q)^m, \quad a = -\frac{q}{1-q}, \quad b = \frac{(m+1)q}{1-q}, \\
p_k &= \binom{m}{k} q^k (1-q)^{m-k}, \quad k = 0, 1, \ldots, m, \\
\mathrm{E}[N] &= mq, \quad \mathrm{Var}[N] = mq(1-q), \\
\hat{q} &= \hat{\mu}/m, \\
P(z) &= [1 + q(z-1)]^m.
\end{aligned}$$

B.2.1.4 Negative binomial—β, r

$$\begin{aligned}
p_0 &= (1+\beta)^{-r}, \quad a = \frac{\beta}{1+\beta}, \quad b = \frac{(r-1)\beta}{1+\beta}, \\
p_k &= \frac{r(r+1)\cdots(r+k-1)\beta^k}{k!(1+\beta)^{r+k}}, \\
\mathrm{E}[N] &= r\beta, \quad \mathrm{Var}[N] = r\beta(1+\beta), \\
\hat{\beta} &= \frac{\hat{\sigma}^2}{\hat{\mu}} - 1, \quad \hat{r} = \frac{\hat{\mu}^2}{\hat{\sigma}^2 - \hat{\mu}}, \\
P(z) &= [1 - \beta(z-1)]^{-r}.
\end{aligned}$$

B.3 THE $(a, b, 1)$ CLASS

To distinguish this class from the $(a, b, 0)$ class, the probabilities are denoted $\Pr(N = k) = p_k^M$ or $\Pr(N = k) = p_k^T$ depending on which subclass is being represented. For this class, p_0^M is arbitrary (i.e., it is a parameter), and then p_1^M or p_1^T is a specified function of the parameters a and b. Subsequent probabilities are obtained recursively as in the $(a, b, 0)$ class: $p_k^M = (a + b/k)p_{k-1}^M$, $k = 2, 3, \ldots$, with the same recursion for p_k^T There are two subclasses of this class. When discussing their members, we often refer to the "corresponding" member of the $(a, b, 0)$ class. This refers to the member of that class with the same values for a and b. The notation p_k will continue to be used for probabilities for the corresponding $(a, b, 0)$ distribution.

B.3.1 The zero-truncated subclass

The members of this class have $p_0^T = 0$, and therefore it need not be estimated. These distributions should only be used when a value of zero is impossible. The first factorial moment is $\mu_{(1)} = (a+b)/[(1-a)(1-p_0)]$, where p_0 is the value for the corresponding member of the $(a, b, 0)$ class. For the logarithmic distribution (which has no corresponding member), $\mu_{(1)} = \beta/\ln(1+\beta)$. Higher factorial moments are obtained recursively with the same formula as with the $(a, b, 0)$ class. The variance is $(a+b)[1-(a+b+1)p_0]/[(1-a)(1-p_0)]^2$. For those members of the subclass that have corresponding $(a, b, 0)$ distributions, $p_k^T = p_k/(1-p_0)$.

B.3.1.1 Zero-truncated Poisson—λ

$$\begin{aligned}
p_1^T &= \frac{\lambda}{e^\lambda - 1}, \quad a = 0, \quad b = \lambda,\\
p_k^T &= \frac{\lambda^k}{k!(e^\lambda - 1)},\\
\mathrm{E}[N] &= \lambda/(1-e^{-\lambda}), \quad \mathrm{Var}[N] = \lambda[1-(\lambda+1)e^{-\lambda}]/(1-e^{-\lambda})^2,\\
\tilde{\lambda} &= \ln(n\hat{\mu}/n_1),\\
P(z) &= \frac{e^{\lambda z}-1}{e^\lambda - 1}.
\end{aligned}$$

B.3.1.2 Zero-truncated geometric—β

$$\begin{aligned}
p_1^T &= \frac{1}{1+\beta}, \quad a = \frac{\beta}{1+\beta}, \quad b = 0,\\
p_k^T &= \frac{\beta^{k-1}}{(1+\beta)^k},\\
\mathrm{E}[N] &= 1+\beta, \quad \mathrm{Var}[N] = \beta(1+\beta),\\
\hat{\beta} &= \hat{\mu} - 1,\\
P(z) &= \frac{[1-\beta(z-1)]^{-1} - (1+\beta)^{-1}}{1-(1+\beta)^{-1}}.
\end{aligned}$$

This is a special case of the zero-truncated negative binomial with $r = 1$.

B.3.1.3 Logarithmic—β

$$\begin{aligned}
p_1^T &= \frac{\beta}{(1+\beta)\ln(1+\beta)}, \quad a = \frac{\beta}{1+\beta}, \quad b = -\frac{\beta}{1+\beta},\\
p_k^T &= \frac{\beta^k}{k(1+\beta)^k \ln(1+\beta)},\\
\mathrm{E}[N] &= \beta/\ln(1+\beta), \quad \mathrm{Var}[N] = \frac{\beta[1+\beta-\beta/\ln(1+\beta)]}{\ln(1+\beta)},\\
\tilde{\beta} &= \frac{n\hat{\mu}}{n_1} - 1 \quad \text{or} \quad \frac{2(\hat{\mu}-1)}{\hat{\mu}},\\
P(z) &= 1 - \frac{\ln[1-\beta(z-1)]}{\ln(1+\beta)}.
\end{aligned}$$

This is a limiting case of the zero-truncated negative binomial as $r \to 0$.

B.3.1.4 Zero-truncated binomial—q, m, $(0 < q < 1, m \text{ an integer})$

$$\begin{aligned}
p_1^T &= \frac{m(1-q)^{m-1}q}{1-(1-q)^m}, \quad a = -\frac{q}{1-q}, \quad b = \frac{(m+1)q}{1-q},\\
p_k^T &= \frac{\binom{m}{k} q^k (1-q)^{m-k}}{1-(1-q)^m}, \quad k = 1, 2, \ldots, m,\\
\mathrm{E}[N] &= \frac{mq}{1-(1-q)^m},\\
\mathrm{Var}[N] &= \frac{mq[(1-q)-(1-q+mq)(1-q)^m]}{[1-(1-q)^m]^2},\\
\tilde{q} &= \frac{\hat{\mu}}{m},\\
P(z) &= \frac{[1+q(z-1)]^m - (1-q)^m}{1-(1-q)^m}.
\end{aligned}$$

B.3.1.5 Zero-truncated negative binomial—$\beta, r, (r > -1, r \neq 0)$

$$\begin{aligned}
p_1^T &= \frac{r\beta}{(1+\beta)^{r+1} - (1+\beta)}, \quad a = \frac{\beta}{1+\beta}, \quad b = \frac{(r-1)\beta}{1+\beta},\\
p_k^T &= \frac{r(r+1)\cdots(r+k-1)}{k![(1+\beta)^r - 1]} \left(\frac{\beta}{1+\beta}\right)^k,\\
\mathrm{E}[N] &= \frac{r\beta}{1-(1+\beta)^{-r}},\\
Var[N] &= \frac{r\beta[(1+\beta)-(1+\beta+r\beta)(1+\beta)^{-r}]}{[1-(1+\beta)^{-r}]^2},\\
\tilde{\beta} &= \frac{\hat{\sigma}^2}{\hat{\mu}} - 1, \quad \tilde{r} = \frac{\hat{\mu}^2}{\hat{\sigma}^2 - \hat{\mu}},\\
P(z) &= \frac{[1-\beta(z-1)]^{-r} - (1+\beta)^{-r}}{1-(1+\beta)^{-r}}.
\end{aligned}$$

This distribution is sometimes called the extended truncated negative binomial distribution because the parameter r can extend below 0.

B.3.2 The zero-modified subclass

A zero-modified distribution is created by starting with a truncated distribution and then placing an arbitrary amount of probability at zero. This probability, p_0^M, is a parameter. The remaining probabilities are adjusted accordingly. Values of p_k^M can be determined from the corresponding zero-truncated distribution as $p_k^M = (1-p_0^M)p_k^T$ or from the corresponding $(a, b, 0)$ distribution as $p_k^M = (1-p_0^M)p_k/(1-p_0)$. The same recursion used for the zero-truncated subclass applies.

The mean is $1-p_0^M$ times the mean for the corresponding zero-truncated distribution. The variance is $1-p_0^M$ times the zero-truncated variance plus $p_0^M(1-p_0^M)$ times the square of the zero-truncated mean. The probability generating function is $P^M(z) = p_0^M + (1-p_0^M)P(z)$, where $P(z)$ is the probability generating function for the corresponding zero-truncated distribution.

The maximum likelihood estimator of p_0^M is always the sample relative frequency at 0.

B.4 THE COMPOUND CLASS

Members of this class are obtained by compounding one distribution with another. That is, let N be a discrete distribution, called the *primary distribution*, and let $M_1, M_2, \ldots$ be i.i.d. with another discrete distribution, called the *secondary distribution*. The compound distribution is $S = M_1 + \cdots + M_N$. The probabilities for the compound distributions are found from

$$p_k = \frac{1}{1-af_0}\sum_{y=1}^{k}(a+by/k)f_y p_{k-y}$$

for $n = 1, 2, \ldots$, where a and b are the usual values for the primary distribution (which must be a member of the $(a, b, 0)$ class) and f_y is p_y for the secondary distribution. The only two primary distributions used here are Poisson (for which $p_0 = \exp[-\lambda(1-f_0)]$) and geometric (for which $p_0 = 1/[1+\beta-\beta f_0]$). Because this information completely describes these distributions, only the names and starting values are given in the following subsections.

The moments can be found from the moments of the individual distributions:

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

The pgf is $P(z) = P_{\text{primary}}[P_{\text{secondary}}(z)]$.

In the following list, the primary distribution is always named first. For the first, second, and fourth distributions, the secondary distribution is the $(a, b, 0)$ class member with that name. For the third and the last three distributions (the Poisson–ETNB and its two special cases), the secondary distribution is the zero-truncated version.

B.4.1 Some compound distributions

B.4.1.1 Poisson–binomial—λ, q, m, $(0 < q < 1$, m an integer)

$$\hat{q} = \frac{\hat{\sigma}^2/\hat{\mu} - 1}{m-1}, \quad \hat{\lambda} = \frac{\hat{\mu}}{m\hat{q}} \quad \text{or} \quad \tilde{q} = 0.5,\ \tilde{\lambda} = \frac{2\hat{\mu}}{m}.$$

B.4.1.2 Poisson–Poisson—λ_1, λ_2

The parameter λ_1 is for the primary Poisson distribution, and λ_2 is for the secondary Poisson distribution. This distribution is also called the *Neyman Type A*.

$$\tilde{\lambda}_1 = \tilde{\lambda}_2 = \sqrt{\hat{\mu}}.$$

B.4.1.3 Geometric–extended truncated negative binomial—β_1, β_2, r $(r > -1)$

The parameter β_1 is for the primary geometric distribution. The last two parameters are for the secondary distribution, noting that for $r = 0$, the secondary distribution is logarithmic. The truncated version is used so that the extension of r is available.

$$\tilde{\beta}_1 = \tilde{\beta}_2 = \sqrt{\hat{\mu}}.$$

B.4.1.4 Geometric–Poisson—β, λ

$$\tilde{\beta} = \tilde{\lambda} = \sqrt{\hat{\mu}}.$$

B.4.1.5 Poisson–extended truncated negative binomial—$\lambda, \beta, (r > -1, r \neq 0)$

When $r = 0$ the secondary distribution is logarithmic, resulting in the negative binomial distribution.

$$\tilde{r} = \frac{\hat{\mu}(K - 3\hat{\sigma}^2 + 2\hat{\mu}) - 2(\hat{\sigma}^2 - \hat{\mu})^2}{\hat{\mu}(K - 3\hat{\sigma}^2 + 2\hat{\mu}) - (\hat{\sigma}^2 - \hat{\mu})^2}, \quad \tilde{\beta} = \frac{\hat{\sigma}^2 - \hat{\mu}}{\hat{\mu}(1+\hat{r})}, \quad \tilde{\lambda} = \frac{\hat{\mu}}{\hat{r}\hat{\beta}},$$

or,

$$\tilde{r} = \frac{\hat{\sigma}^2 n_1/n - \hat{\mu}^2 n_0/n}{(\hat{\sigma}^2 - \hat{\mu}^2)(n_0/n)\ln(n_0/n) - \hat{\mu}(\hat{\mu}n_0/n - n_1/n)},$$

$$\tilde{\beta} = \frac{\hat{\sigma}^2 - \hat{\mu}}{\hat{\mu}(1+\hat{r})}, \quad \tilde{\lambda} = \frac{\hat{\mu}}{\hat{r}\hat{\beta}}$$

where

$$K = \frac{1}{n}\sum_{k=0}^{\infty} k^3 n_k - 3\hat{\mu}\frac{1}{n}\sum_{k=0}^{\infty} k^2 n_k + 2\hat{\mu}^3.$$

This distribution is also called the *generalized Poisson–Pascal*.

B.4.1.6 Polya–Aeppli—λ, β

$$\hat{\beta} = \frac{\hat{\sigma}^2 - \hat{\mu}}{2\hat{\mu}}, \quad \hat{\lambda} = \frac{\hat{\mu}}{1+\hat{\beta}}.$$

This is a special case of the Poisson–extended truncated negative binomial with $r = 1$. It is actually a Poisson–truncated geometric.

B.4.1.7 Poisson–inverse Gaussian—λ, β

$$\tilde{\lambda} = -\ln(n_0/n), \quad \tilde{\beta} = \frac{4(\hat{\mu} - \tilde{\lambda})}{\hat{\mu}}.$$

This is a special case of the Poisson–extended truncated negative binomial with $r = -0.5$.

Table B.1 Hierarchy of discrete distributions.

Distribution	Is a special case of	Is a limiting case of
Poisson	ZM Poisson	Negative binomial, Poisson–binomial, Poisson–inv. Gaussian, Polya–Aeppli, Neyman–A
ZT Poisson	ZM Poisson	ZT negative binomial
ZM Poisson		ZM negative binomial
Geometric	Negative binomial ZM geometric	Geometric–Poisson
ZT geometric	ZT negative binomial	
ZM geometric	ZM negative binomial	
Logarithmic		ZT negative binomial
ZM logarithmic		ZM negative binomial
Binomial	ZM binomial	
Negative binomial	ZM negative binomial	Poisson–ETNB
Poisson–inverse Gaussian	Poisson–ETNB	
Polya–Aeppli	Poisson–ETNB	
Neyman–A		Poisson–ETNB

B.5 A HIERARCHY OF DISCRETE DISTRIBUTIONS

Table B.1 indicates which distributions are special or limiting cases of others. For the special cases, one parameter is set equal to a constant to create the special case. For the limiting cases, two parameters go to infinity or zero in some special way.

Appendix C
Frequency and Severity Relationships

Let N^L be the number of losses random variable and let X be the severity random variable. If there is a deductible of d imposed, there are two ways to modify X. One is to create Y^L, the amount paid per loss:

$$Y^L = \begin{cases} 0, & X \le d, \\ X - d, & X > d. \end{cases}$$

In this case, the appropriate frequency distribution continues to be N^L.

An alternative approach is to create Y^P, the amount paid per payment:

$$Y^P = \begin{cases} \text{undefined}, & X \le d, \\ X - d, & X > d. \end{cases}$$

In this case, the frequency random variable must be altered to reflect the number of payments. Let this variable be N^P. Assume that for each loss the probability is $v = 1 - F_X(d)$ that a payment will result. Further assume that the incidence of making a payment is independent of the number of losses. Then $N^P = L_1 + L_2 + \cdots + L_N$, where L_j is 0 with probability $1 - v$ and is 1 with probability v. Probability generating functions yield the relationships in Table C.1.

Loss Models: From Data to Decisions, 3rd. ed.
By Stuart A. Klugman, Harry H. Panjer, Gordon E. Willmot

Table C.1 Parameter adjustments.

N^L	Parameters for N^P
Poisson	$\lambda^* = v\lambda$
ZM Poisson	$p_0^{M*} = \dfrac{p_0^M - e^{-\lambda} + e^{-v\lambda} - p_0^M e^{-v\lambda}}{1 - e^{-\lambda}}$, $\lambda^* = v\lambda$
Binomial	$q^* = vq$
ZM binomial	$p_0^{M*} = \dfrac{p_0^M - (1-q)^m + (1-vq)^m - p_0^M(1-vq)^m}{1-(1-q)^m}$ $q^* = vq$
Negative binomial	$\beta^* = v\beta,\ r^* = r$
ZM neg. binomial	$p_0^{M*} = \dfrac{p_0^M - (1+\beta)^{-r} + (1+v\beta)^{-r} - p_0^M(1+v\beta)^{-r}}{1-(1+\beta)^{-r}}$ $\beta^* = v\beta,\ r^* = r$
ZM logarithmic	$p_0^{M*} = 1 - (1-p_0^M)\ln(1+v\beta)/\ln(1+\beta)$ $\beta^* = v\beta$

The geometric distribution is not presented as it is a special case of the negative binomial with $r = 1$. For zero-truncated distributions, the same formulas are still used as the distribution for N^P will now be zero modified. For compound distributions, modify only the secondary distribution. For ETNB, secondary distributions the parameter for the primary distribution is multiplied by $1 - p_0^{M*}$ as obtained in Table C.1, while the secondary distribution remains zero truncated (however, $\beta^* = v\beta$).

There are occasions in which frequency data are collected that provide a model for N^P. There would have to have been a deductible d in place and therefore v is available. It is possible to recover the distribution for N^L, although there is no guarantee that reversing the process will produce a legitimate probability distribution. The solutions are the same as in Table C.1, only now $v = 1/[1 - F_X(d)]$.

Now suppose the current frequency model is N^d, which is appropriate for a deductible of d. Also suppose the deductible is to be changed to d^*. The new frequency for payments is N^{d^*} and is of the same type. Then use Table C.1 with $v = [1 - F_X(d^*)]/[1 - F_X(d)]$.

Appendix D
The Recursive Formula

The recursive formula is (where the frequency distribution is a member of the $(a, b, 1)$ class),

$$f_S(x) = \frac{[p_1 - (a+b)p_0]f_X(x) + \sum_{y=1}^{x \wedge m}\left(a + \frac{by}{x}\right) f_X(y) f_S(x-y)}{1 - a f_X(0)},$$

where $f_S(x) = \Pr(S = x)$, $x = 0, 1, 2, \ldots$, $f_X(x) = \Pr(X = x)$, $x = 0, 1, 2, \ldots$, $p_0 = \Pr(N = 0)$, and $p_1 = \Pr(N = 1)$. Note that the severity distribution (X) must place probability on nonnegative integers. The formula must be initialized with the value of $f_S(0)$. These values are given in Table D.1. It should be noted that, if N is a member of the $(a, b, 0)$ class, $p_1 - (a+b)p_0 = 0$, and so the first term will vanish. If N is a member of the compound class, the recursion must be run twice. The first pass uses the secondary distribution for p_0, p_1, a, and b. The second pass uses the output from the first pass as $f_X(x)$ and uses the primary distribution for p_0, p_1, a, and b.

Loss Models: From Data to Decisions, 3rd. ed.
By Stuart A. Klugman, Harry H. Panjer, Gordon E. Willmot

Table D.1 Starting values ($f_S(0)$) for recursions.

Distribution	$f_S(0)$
Poisson	$\exp[\lambda(f_0 - 1)]$
Geometric	$[1 + \beta(1 - f_0)]^{-1}$
Binomial	$[1 + q(f_0 - 1)]^m$
Negative binomial	$[1 + \beta(1 - f_0)]^{-r}$
ZM Poisson	$p_0^M + (1 - p_0^M)\dfrac{\exp(\lambda f_0) - 1}{\exp(\lambda) - 1}$
ZM geometric	$p_0^M + (1 - p_0^M)\dfrac{f_0}{1 + \beta(1 - f_0)}$
ZM binomial	$p_0^M + (1 - p_0^M)\dfrac{[1 + q(f_0 - 1)]^m - (1 - q)^m}{1 - (1 - q)^m}$
ZM negative binomial	$p_0^M + (1 - p_0^M)\dfrac{[1 + \beta(1 - f_0)]^{-r} - (1 + \beta)^{-r}}{1 - (1 + \beta)^{-r}}$
ZM logarithmic	$p_0^M + (1 - p_0^M)\left\{1 - \dfrac{\ln[1 + \beta(1 - f_0)]}{\ln(1 + \beta)}\right\}$

Appendix E
Discretization of the Severity Distribution

There are two relatively simple ways to discretize the severity distribution. One is the method of rounding and the other is a mean-preserving method.

E.1 THE METHOD OF ROUNDING

This method has two features: All probabilities are positive and the probabilities add to 1. Let h be the span and let Y be the discretized version of X. If there are no modifications, then

$$\begin{aligned} f_j &= \Pr(Y = jh) = \Pr\left[\left(j - \tfrac{1}{2}\right)h \le X < \left(j + \tfrac{1}{2}\right)h\right] \\ &= F_X\left[\left(j + \tfrac{1}{2}\right)h\right] - F_X\left[\left(j - \tfrac{1}{2}\right)h\right]. \end{aligned}$$

The recursive formula is then used with $f_X(j) = f_j$. Suppose a deductible of d, limit of u, and coinsurance of α are to be applied. If the modifications are to be applied before the

Loss Models: From Data to Decisions, 3rd. ed.
By Stuart A. Klugman, Harry H. Panjer, Gordon E. Willmot

discretization, then

$$\begin{aligned} g_0 &= \frac{F_X(d + h/2) - F_X(d)}{1 - F_X(d)}, \\ g_j &= \frac{F_X[d + (j + 1/2)h] - F_X[d + (j - 1/2)h]}{1 - F_X(d)}, \\ & \qquad j = 1, \ldots, \frac{u-d}{h} - 1, \\ g_{(u-d)/h} &= \frac{1 - F_X(u - h/2)}{1 - F_X(d)}, \end{aligned}$$

where $g_j = \Pr(Z = j\alpha h)$ and Z is the modified distribution. This method does not require that the limits be multiples of h but does require that $u - d$ be a multiple of h. This method gives the probabilities of payments per payment.

Finally, if there is truncation from above at u, change all denominators to $F_X(u) - F_X(d)$ and also change the numerator of $g_{(u-d)/h}$ to $F_X(u) - F_X(u - h/2)$.

E.2 MEAN PRESERVING

This method ensures that the discretized distribution has the same mean as the original severity distribution. With no modifications, the discretization is

$$\begin{aligned} f_0 &= 1 - \frac{\mathrm{E}[X \wedge h]}{h}, \\ f_j &= \frac{2\mathrm{E}[X \wedge jh] - \mathrm{E}[X \wedge (j-1)h] - \mathrm{E}[X \wedge (j+1)h]}{h}, \quad j = 1, 2, \ldots. \end{aligned}$$

For the modified distribution,

$$\begin{aligned} g_0 &= 1 - \frac{\mathrm{E}[X \wedge d + h] - \mathrm{E}[X \wedge d]}{h[1 - F_X(d)]}, \\ g_j &= \frac{2\mathrm{E}[X \wedge d + jh] - \mathrm{E}[X \wedge d + (j-1)h] - \mathrm{E}[X \wedge d + (j+1)h]}{h[1 - F_X(d)]}, \\ & \qquad j = 1, \ldots, \frac{u-d}{h} - 1, \\ g_{(u-d)/h} &= \frac{\mathrm{E}[X \wedge u] - \mathrm{E}[X \wedge u - h]}{h[1 - F_X(d)]}. \end{aligned}$$

To incorporate truncation from above, change the denominators to

$$h[F_X(u) - F_X(d)]$$

and subtract $h[1 - F_X(u)]$ from the numerators of each of g_0 and $g_{(u-d)/h}$.

E.3 UNDISCRETIZATION OF A DISCRETIZED DISTRIBUTION

Assume we have $g_0 = \Pr(S = 0)$, the true probability that the random variable is zero. Let $p_j = \Pr(S^* = jh)$, where S^* is a discretized distribution and h is the span. The following

Compound distributions

$$S = \sum_{i=0}^{N} X_i$$

with $X_0 \equiv 0$, where N denotes the number of claims in the time period and X_i denotes the amount of the $i-th$ claim. We assume that

- $\{X_i\}_{i=1}^{\infty}$ is a sequence of independent and identically distributed random variables.
- N is independent of this sequence.

Let $M_S(t) = E(\exp\{tS\})$. Then $M_S(t) = M_N(\log M_X(t))$ whenever these moment generating functions exist. Also

$$E(S) = E(N)E(X_i),$$

$$V(S) = E(N)V(X_i) + V(N)E(X_i)^2$$

and

$$\mu_3[S] = \mu_3[N]E^3[X_i] + 3\mathrm{Var}[N]\mathrm{E}[X_i]\mathrm{Var}[X_i] + \mathrm{E}[N]\mu_3(X_i).$$

The NP approximation

Let $F_S(x)$ be the distribution function of S and let $F_Z(x)$ be the distribution function of $Z = (S - \mu_S)/\sigma_S$. Then

$$F_Z\left(z + \frac{\gamma_S}{6}(z^2 - 1)\right) \approx \Phi(z),$$

which is equivalent to

$$F_S(x) \approx \Phi\left(-\frac{3}{\gamma_S} + \sqrt{\frac{9}{\gamma_S^2} + 1 + \frac{6}{\gamma_S}\frac{x - \mu_S}{\sigma_S}}\right),$$

where Φ is the distribution function of the standardized normal random variable, and μ_S, σ_S and γ_S are, respectively, the expected value, standard deviation and skewness coefficient of S.