Excerpts from the Appendices to *Loss Models: From Data to Decisions, 3rd edition*

August 27, 2009

Appendix A

An Inventory of Continuous Distributions

A.1 Introduction

The incomplete gamma function is given by

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

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

Also, define

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

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

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

The incomplete beta function is given by

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

A.2 Transformed beta family

A.2.2 Three-parameter distributions

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

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

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

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

A.2.2.3 Inverse Burr (Dagum)—τ, θ, γ

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

A.2.3 Two-parameter distributions

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

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

A.2.3.2 Inverse Pareto—τ, θ

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

A.2.3.3 Loglogistic (Fisk)—γ, θ

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

A.2.3.4 Paralogistic—α, θ

This is a Burr distribution with $\gamma = \alpha$.

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

A.2.3.5 Inverse paralogistic—τ, θ

This is an inverse Burr distribution with $\gamma = \tau$.

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

A.3 Transformed gamma family

A.3.2 Two-parameter distributions

A.3.2.1 Gamma—α, θ

$$
\begin{aligned}
f(x) &= \frac{(x/\theta)^\alpha e^{-x/\theta}}{x \Gamma(\alpha)} \qquad F(x) = \Gamma(\alpha; x/\theta) \\
M(t) &= (1 - \theta t)^{-\alpha}, \quad t < 1/\theta \qquad \mathrm{E}[X^k] = \frac{\theta^k \Gamma(\alpha + k)}{\Gamma(\alpha)}, \quad k > -\alpha \\
\mathrm{E}[X^k] &= \theta^k (\alpha + k - 1) \cdots \alpha, \quad \text{if } k \text{ is an integer}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k \Gamma(\alpha + k)}{\Gamma(\alpha)} \Gamma(\alpha + k; x/\theta) + x^k [1 - \Gamma(\alpha; x/\theta)], \quad k > -\alpha \\
&= \alpha(\alpha+1) \cdots (\alpha + k - 1) \theta^k \Gamma(\alpha + k; x/\theta) + x^k [1 - \Gamma(\alpha; x/\theta)], \quad k \text{ an integer} \\
\text{mode} &= \theta(\alpha - 1), \quad \alpha > 1, \text{ else } 0
\end{aligned}
$$

A.3.2.2 Inverse gamma (Vinci)—α, θ

$$
\begin{aligned}
f(x) &= \frac{(\theta/x)^\alpha e^{-\theta/x}}{x\Gamma(\alpha)} \qquad F(x) = 1 - \Gamma(\alpha;\theta/x) \\
\mathrm{E}[X^k] &= \frac{\theta^k\Gamma(\alpha-k)}{\Gamma(\alpha)}, \quad k < \alpha \qquad \mathrm{E}[X^k] = \frac{\theta^k}{(\alpha-1)\cdots(\alpha-k)}, \quad \text{if } k \text{ is an integer} \\
\mathrm{E}[(X \wedge x)^k] &= \frac{\theta^k\Gamma(\alpha-k)}{\Gamma(\alpha)}[1 - \Gamma(\alpha-k;\theta/x)] + x^k\Gamma(\alpha;\theta/x) \\
&= \frac{\theta^k\Gamma(\alpha-k)}{\Gamma(\alpha)} G(\alpha-k;\theta/x) + x^k\Gamma(\alpha;\theta/x), \text{ all } k \\
\text{mode} &= \theta/(\alpha+1)
\end{aligned}
$$

A.3.2.3 Weibull—θ, τ

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

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

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

A.3.3 One-parameter distributions

A.3.3.1 Exponential—θ

$$
\begin{aligned}
f(x) &= \frac{e^{-x/\theta}}{\theta} \qquad F(x) = 1 - e^{-x/\theta} \\
M(t) &= (1-\theta t)^{-1} \qquad \mathrm{E}[X^k] = \theta^k \Gamma(k+1), \quad k > -1 \\
\mathrm{E}[X^k] &= \theta^k k!, \quad \text{if } k \text{ is an integer} \\
\mathrm{VaR}_p(X) &= -\theta \ln(1-p) \\
\mathrm{TVaR}_p(X) &= -\theta \ln(1-p) + \theta \\
\mathrm{E}[X \wedge x] &= \theta(1 - e^{-x/\theta}) \\
\mathrm{E}[(X \wedge x)^k] &= \theta^k \Gamma(k+1)\Gamma(k+1; x/\theta) + x^k e^{-x/\theta}, \quad k > -1 \\
&= \theta^k k! \Gamma(k+1; x/\theta) + x^k e^{-x/\theta}, \quad k \text{ an integer} \\
\text{mode} &= 0
\end{aligned}
$$

A.3.3.2 Inverse exponential—θ

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

A.5 Other distributions

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

$$
\begin{aligned}
f(x) &= \frac{1}{x\sigma\sqrt{2\pi}} \exp(-z^2/2) = \phi(z)/(\sigma x), \quad z = \frac{\ln x - \mu}{\sigma} \qquad F(x) = \Phi(z) \\
\mathrm{E}[X^k] &= \exp(k\mu + k^2\sigma^2/2) \\
\mathrm{E}[(X \wedge x)^k] &= \exp(k\mu + k^2\sigma^2/2)\Phi\left(\frac{\ln x - \mu - k\sigma^2}{\sigma}\right) + x^k[1 - F(x)] \\
\text{mode} &= \exp(\mu - \sigma^2)
\end{aligned}
$$

A.5.1.2 Inverse Gaussian—μ, θ

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

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.

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

$$
F(x) = \begin{cases} \frac{1}{2}\beta\left[\frac{r}{2}, \frac{1}{2}; \frac{r}{r+\left(\frac{\ln x-\mu}{\sigma}\right)^2}\right], & 0 < x \le e^{\mu}, \\ 1-\frac{1}{2}\beta\left[\frac{r}{2}, \frac{1}{2}; \frac{r}{r+\left(\frac{\ln x-\mu}{\sigma}\right)^2}\right], & x \ge e^{\mu}. \end{cases}
$$

A.5.1.4 Single-parameter Pareto—α, θ

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

Note: Although there appears 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

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

Appendix B

An Inventory of Discrete Distributions

B.1 Introduction

The 16 models 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 which are presented are not intended to be useful estimators but rather for providing 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]:

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

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 and/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

B.2.1.1 Poisson—λ

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

B.2.1.2 Geometric—β

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

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

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

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

B.2.1.4 Negative binomial—β, r

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

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 (that is, 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 sub-classes 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 which 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—$q, m, (0 < q < 1,\ m$ 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.