Severity Distributions

Fall 2017

Outline

Important Severity Distributions

Three important loss severity distributions:

- **Gamma**
  - Fits medium tail lines like physical damage auto and homeowners well
  - Member of the "exponential family of distributions". This means that it is easy to incorporate rating variables into the distribution via generalized linear modeling (GLMs)
- **Pareto**
  - Fits longer tail lines like injury liability in auto and workers' compensation well
  - Simple to work with analytically (hence can provide intuition as we develop theory and explain the theory to others)
- **GB2 - Generalized Beta of the Second Kind**
  - A four parameter distribution family, complex
  - Yet, many severity distributions can be expressed as a special case of this distribution (good for programming)
  - Some applications have been fit well by the GB2 where others do not seem to work

Gamma Distribution

- The gamma distribution has two parameters, α and θ
- The probability density function is 0 for $x \leq 0$ and for $x > 0$

$$f(x) \quad = \quad \frac{\left(\frac{x}{\theta}\right)^{\alpha} e^{-x/\theta}}{x\Gamma(\alpha)} = \frac{1}{\theta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta}$$

- If $\alpha = 1$, the gamma reduces to the familiar *exponential* distribution
- The function $\Gamma(\cdot)$ is known as the *gamma function*, defined as

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

- Some important facts about the gamma function:
    - For a positive integer n, $\Gamma(n) = (n-1)!$
    - For more general arguments, one needs to rely on numerical integration to evaluate $\Gamma(\cdot)$. The two main exceptions are:
        - For any $\alpha > 0$, $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$
        - $\Gamma(0.5) = \sqrt{\pi}$

Thus, for example, $\Gamma(2.5) = 1.5\Gamma(1.5) = 1.5(0.5)\Gamma(0.5) = \frac{3}{4}\sqrt{\pi}$

Exercise

Example. Suppose that $X \sim gamma(\alpha = 2, \theta = 100)$, that is, the random variable X has a gamma distribution with parameters $\alpha = 2$ and $\theta = 100$

Determine $\Pr(X \leq 60)$

Gamma Moments

Use the gamma function to calculate moments of a gamma distribution

- Define the kth *raw moment* to be

$$\mu'_k = \mathrm{E}\, X^k = \int_0^\infty x^k f(x)dx$$

- Using a change of variable, $t = x/\theta$, we have

$$\begin{aligned}\mu'_k &= \frac{1}{\theta^\alpha \Gamma(\alpha)} \int_0^\infty x^{\alpha+k-1} \exp(-x/\theta)\, dx \\ &= \frac{\theta^{\alpha+k}}{\theta^\alpha \Gamma(\alpha)} \int_0^\infty t^{\alpha+k-1} \exp(-t) dt \\ &= \frac{\theta^k}{\Gamma(\alpha)} \Gamma(\alpha+k).\end{aligned}$$

- With $k = 1$, we have $\mu = \mu'_1 = \theta \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} = \alpha\theta$
- Check that:
  - $\mu'_2 = \theta^2 \alpha(\alpha+1), \qquad \mathrm{Var}(X) = \theta^2\alpha, \qquad \mu'_k = \theta^k(\alpha+k-1)\cdots\alpha$

Gamma Moment Generating Function

- The gamma moment generating function (mgf) is

$$\begin{aligned} M(t) = \mathrm{E}\ e^{tX} &= \frac{1}{\theta^{\alpha}\Gamma(\alpha)} \int_0^{\infty} x^{\alpha-1} e^{tx} e^{-x/\theta} dx \\ &= \frac{1}{\theta^{\alpha}\Gamma(\alpha)} \int_0^{\infty} x^{\alpha-1} \exp\left\{ -x \left(\frac{1-\theta t}{\theta} \right) \right\} dx \\ &= \frac{1}{\theta^{\alpha}\Gamma(\alpha)} \left(\frac{\theta}{1-\theta t} \right)^{\alpha} \Gamma(\alpha) \\ &= (1-\theta t)^{-\alpha}. \end{aligned}$$

- To check this result, first note that $M(0) = (1-\theta 0)^{-\alpha} = 1$, as anticipated. Next, taking derivatives, we have

$$M'(t) = \frac{\partial}{\partial t} M(t) = -\alpha(1-\theta t)^{-\alpha-1}(-\theta) = \alpha\theta(1-\theta t)^{-\alpha-1}$$

- Evaluating this at 0 yields $M'(0) = \alpha\theta = \mu$, as anticipated

Pareto Distribution

- The Pareto with parameters α and θ has probability density function:

$$\mathrm{f}(x) = \frac{\alpha\theta^{\alpha}}{(x+\theta)^{\alpha+1}}$$

  and moments

$$\mu_k' = \frac{\theta^k k!}{(\alpha-1)\cdots(\alpha-k)}$$

- Because moments are not finite when $k \geq \alpha$, the moment generating function is not well-defined
- Unlike the gamma, there is a simple expression for the distribution function

$$\mathrm{F}(x) \quad = \quad \int_0^x \mathrm{f}(y)dy = 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha}$$

- This means that it is easy to compute quantiles
  - For example, find x so that 0.95 = F(x) (the 95th percentile)
  - Easy calculations show that this is $\theta\left[(0.05)^{-1/\alpha} - 1\right]$
  - In general, the pth percentile/quantile is $\theta\left[(1-p)^{-1/\alpha} - 1\right]$

GB2 - Generalized Beta of the Second Kind

- In KPW, the GB2 is known as a *transformed beta distribution*
- The pdf (KPW, Appendix A2.1.1) is

$$f(x) = \frac{\Gamma(\alpha+\tau)}{\Gamma(\alpha)\Gamma(\tau)} \frac{\gamma(x/\theta)^{\gamma\tau}}{x\left[1+(x/\theta)^{\gamma}\right]^{\alpha+\tau}}$$

with moments

$$\mathrm{E}\, X^k = \theta^k \frac{\Gamma(\tau+\frac{k}{\gamma})\Gamma(\alpha-\frac{k}{\gamma})}{\Gamma(\alpha)\Gamma(\tau)}.$$

- In the text by Frees on Regression, the GB2 distribution is cited but with the parameters $\alpha_1=\alpha$, $\alpha_2=\tau$, $\gamma=1/\sigma$, $\theta=e^{\mu}$

GB2 Special Cases

The GB2 a four parameter family of distributions that captures many other distributions, either as special cases or as limiting results:

- *Special Case: Burr Distribution*. Use the GB2 distribution with $\tau = 1$
- *Special Case: Pareto Distribution*. Use the GB2 distribution with $\gamma = \tau = 1$
- *Limiting Case: Generalized Gamma Distribution*

  Replace θ by $\theta\tau^{1/\gamma}$

  Then, one can show that

$$\lim_{\tau \rightarrow \infty} f_{GB2}(x; \theta\tau^{1/\gamma}, \alpha, \tau, \gamma) = f(x),$$

  the pdf of a *generalized gamma*. In KPW, Appendix A.3.1, p.673, the generalized gamma is called a *transformed gamma*

Creating Distributions Using Transformations

There are many distributions available to the analyst

- In this section, we consider distributions that are created by transforming the random variable of a distribution. Specifically:
    - Multiplication by a constant ($Y = cX$)
    - Raising to a power ($Y = X^{\tau}$)
    - Exponentiation ($Y = e^X$)
- In the next section, we consider ways of combining distributions to form a distribution of interest. Specifically:
    - Mixing
    - Splicing

Multiplication by a Constant

- Multiplying a random variable by a positive constant is a simple type of transformation
- It is also easy to interpret:
  - Think of X as this year's losses and assume that we have an 8% inflation rate. Then, we can model next year's losses as $Y = 1.08X$
  - We also want to readily go from dollars to thousands of dollars ($c = 1/1000$) or from dollars to Euros (or swapping any set of currencies)
- More generally, let $Y = cX$ and use

$$
\begin{aligned}
F_Y(y) &= \Pr(Y \leq y) = \Pr\left(X \leq \frac{y}{c}\right) = F_X\left(\frac{y}{c}\right) \\
f_Y(y) &= \frac{1}{c} f_X\left(\frac{y}{c}\right)
\end{aligned}
$$

Scale Distributions

- In a location-scale distribution, the transformed variable $Y = c(X - a)$ has a distribution from the same family as the random variable X
  - Here, a and $c > 0$ are constants
  - The normal distribution is the usual example
- In a scale distribution, the transformed variable $Y = cX$ has a distribution from the same family as the random variable X
  - Many loss distributions are scale distributions
  - Typically, one uses θ as the scale parameter. If X comes from a distribution with parameter θ, then $Y = cX$ has the same distribution with scale parameter $\theta^* = c\theta$
- *Example: Special Case - Gamma Distribution*. Suppose that X has a gamma distribution with $\alpha = 4$ and $\theta = 10,000$
  - The mean is $\alpha\theta = 40,000$ and the standard deviation is $\theta\sqrt{\alpha} = 20,000$
  - Suppose that $Y = X/1000$
    - This has mean 40 and standard deviation 20
    - Further Y has a gamma distribution with parameters $\alpha = 4$ and $\theta^* = \theta/1000 = 10$

Raising to a Power

- Another type of transformation involves raising the random variable to a power, say, τ
- Consider the transformed random variable $Y = X^{\tau}$. We examine three cases:

$$\begin{array}{rl} \tau > 0 & \text{transformed} \\ \tau = -1 & \text{inverse} \\ \tau < 0 & \text{inverse transformed} \end{array}$$

- *Special Case: Exponential Distribution*. Suppose that X has an exponential distribution with parameter θ^* and consider $Y = 1/X$
  - The distribution function of Y is

$$\Pr(Y \le y) = \Pr(\frac{1}{X} \le y) = \Pr(\frac{1}{y} \le X) = \exp\left(-\frac{1}{y\theta^*}\right).$$

  - Now, define a new parameter $\theta = \frac{1}{\theta^*}$. With this notation,

$$\Pr(Y \le y) = \exp\left(-\frac{\theta}{y}\right).$$

  - This distribution is known as an *inverse exponential distribution* with parameter θ. See the appendix of KPW

Exponential to get a Weibull

Example: Transforming an Exponential to get a Weibull.

- Start with $X \sim$ exponential distribution with parameter 1. Define the transformed random variable:

$$Y = \theta X^{1/\tau}$$

- This has distribution

$$\begin{aligned} F_Y(y) &= \Pr(Y \leq y) \\ &= \Pr(X^{1/\tau} \leq \frac{y}{\theta}) = \Pr(X \leq (\frac{y}{\theta})^\tau) \\ &= 1 - \exp\left(-(\frac{y}{\theta})^\tau\right), \end{aligned}$$

  known as a *Weibull distribution*

- This result will be handy if you want to *simulate* outcomes from a Weibull distribution in Excel (exponential simulation is easy, Weibull is not available)

Transforming the Pareto Distribution

This is from KPW Exercise 5.3. We assume that $X \sim$ Pareto with parameters (α, θ) and consider the transformed variable $Y = X^{1/\tau}$. We wish to determine the df of Y when τ is positive, equal to -1, and negative

Solution. Begin by recalling the df of the Pareto

$$\mathrm{F}_X(x) \quad = \quad 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha}.$$

Case 1. Assume $\tau > 0$. Then,

$$\begin{aligned}
\mathrm{F}_Y(y) &= \Pr(X^{1/\tau} \le y) = \Pr(X \le y^{\tau}) \\
&= \mathrm{F}_X(y^{\tau}) \\
&= 1 - \left(\frac{\theta}{y^{\tau}+\theta}\right)^{\alpha}.
\end{aligned}$$

Now, define the new parameter $\theta^* = \theta^{1/\tau}$ so that $\theta^{*\tau} = \theta$. With this notation, we have

$$\mathrm{F}_Y(y) \quad = \quad 1 - \left(\frac{\theta^{*\tau}}{y^{\tau}+\theta^{*\tau}}\right)^{\alpha}.$$

This is known as a *Burr distribution* with parameters $(\alpha, \theta^*, \tau = \gamma)$

Exponentiation

- Another type of transformation involves exponentiating a random variable so that $Y = \exp(X)$
- The main example of this is the normal distribution. If $X \sim$ normal, then $Y = e^X \sim$ a *lognormal distribution*
- We can develop the distribution of the new random variable through the relation with the df

$$F_Y(y) = \Pr(\exp(X) \leq y) = \Pr(X \leq \ln y) = F_X(\ln y)$$

and the pdf

$$f_Y(y) = \frac{1}{y} f_X(\ln y).$$

- Remark. This provides a way to simulate a Pareto distribution, by first simulating an exponential random variable and then transforming it

Motivation for Mixing

- In a mixture distribution, the outcome (random variable) can be thought of as a random draw from a population of outcomes
- *Example: Pareto Distribution*. Consider the Pareto distribution with survival function

$$S(x) = \left(\frac{\theta}{x+\theta}\right)^{\alpha},$$

and mean $\mathrm{E}\, X = \frac{\theta}{\alpha-1}$. Let us think about two types of populations:

$X_1 \sim$ Pareto(α_1, θ_1) "Good Driver"
$X_2 \sim$ Pareto(α_2, θ_2) "Bad Driver"

- Suppose that with probability a we draw the loss from a good driver, X_1, and with probability $1-a$ we draw the loss from a bad driver, X_2:

$$Y = \begin{cases} X_1 & \text{with probability } a \\ X_2 & \text{with probability } 1-a \end{cases}$$

Our interest is in the distribution of Y

Mixing Moments

- To begin, focus on the mean. Using the law of total expectations:

$$\mathrm{E}\ Y = a\mathrm{E}\ X_1 + (1-a)\mathrm{E}\ X_2 = a\frac{\theta_1}{\alpha_1 - 1} + (1-a)\frac{\theta_2}{\alpha_2 - 1}.$$

We can also write

$$Y^2 = \begin{cases} X_1^2 & \text{with probability } a \\ X_2^2 & \text{with probability } 1-a \end{cases}$$

Thus, we have

$$\mathrm{E}\ Y^2 = a\mathrm{E}\ X_1^2 + (1-a)\mathrm{E}\ X_2^2.$$

The same argument holds for any moment

Mixing Distribution Function

- For the distribution function, we have

$$\begin{aligned}
\Pr(Y \le y) &= \Pr(Y \le y, \text{Good Driver}) + \Pr(Y \le y, \text{Bad Driver}) \\
&= \Pr(X_1 \le y, \text{Good Driver}) + \Pr(X_2 \le y, \text{Bad Driver}) \\
&= \Pr(X_1 \le y) \Pr(\text{Good Driver}) + \Pr(X_2 \le y) \Pr(\text{Bad Driver}) \\
&= aF_{X_1}(y) + (1-a)F_{X_2}(y)
\end{aligned}$$

Example from Exam M Spring 05 #34. Suppose that a = 0.8, and $X_1 \sim$ Pareto$(\alpha = 2, \theta = 100), X_2 \sim$ Pareto$(\alpha = 4, \theta = 3000)$

Determine $\Pr(Y \le 200)$

Finite Mixture Distributions

- *Definition.* Let $X_1, \ldots, X_k$ be random variables and define

$$Y = \begin{cases} X_1 & \textit{with probability } a_1 \\ \vdots & \vdots \\ X_k & \textit{with probability } a_k \end{cases}$$

  Here, $a_j > 0$ and $a_1 + \cdots + a_k = 1$. Then, Y is a k-point mixture random variable. The df is

$$F_Y(y) = a_1 F_{X_1}(y) + \cdots + a_k F_{X_k}(y)$$

  with mean

$$\mathrm{E}\ Y = a_1 \mathrm{E}\ X_1 + \cdots + a_k \mathrm{E}\ X_k.$$

- If k is unknown (but not random), then this a *variable component mixture distribution*
- We can always select one or more of the underlying X_j variables to be degenerate (that is, equal to a number with probability one). In this way, we can use the finite mixture framework to create discrete and mixed distributions

Exponential Example

Example. Suppose, for a fixed parameter θ, that $X|\theta \sim$ exponential with parameter θ. Thus,

$$\Pr(X \leq x|\theta) = 1 - e^{-x/\theta}.$$

Now, think of two populations, each having an exponential distribution, but with different parameter values. For concreteness, assume

$$\Theta = \begin{cases} 10 & \text{with prob } \alpha \\ 200 & \text{with prob } (1-\alpha) \end{cases}$$

As with the x's, we use an upper case Θ for a random variable and a lower case θ for a realization of the random variable

Exponential Example II

- Using the discrete mixing framework, we may write the df of X as

$$\Pr(X \leq x) = \alpha(1 - e^{-x/10}) + (1 - \alpha)(1 - e^{-x/200}).$$

- More generally, we can consider k populations, each with the same form of the distribution function $\mathrm{F}(\cdot|\theta)$ and allow θ to vary as

$$\Theta = \begin{cases} \theta_1 & \text{prob } \alpha_1 \\ \theta_2 & \text{prob } \alpha_2 \\ \vdots & \vdots \\ \theta_K & \text{prob } \alpha_K \end{cases}$$

This is our *finite mixture* distribution

Continuous Mixtures

- Extend this idea by thinking about an infinite number of populations, each with a conditional distribution function that has the same structure $\mathrm{F}(\cdot|\theta)$ (e.g., exponential) but with a parameter θ that accounts for population differences
- Assume that the random variable Θ has pdf $f_\Theta(\theta)$
- Then, the df is:

$$\begin{aligned} \mathrm{F}_X(x) = \Pr(X \le x) &= E_\Theta \Pr(X \le x|\Theta) \\ &= \int \Pr(X \le x|\theta) f_\Theta(\theta) d\theta = \int \mathrm{F}(x|\theta) f_\Theta(\theta) d\theta \end{aligned}$$

- The pdf is:

$$f_X(x) = \int f_{x|\theta}(x) f_\Theta(\theta) d\theta$$

Special Case: Gamma Mixtures of Exponentials

- *Special case: Gamma Mixtures of Exponentials*. Suppose that each population has an exponential distribution with parameter $1/\theta$, that is, $X|\theta \sim$ exponential$(\frac{1}{\theta})$:

$$f_{X|\theta}(x) = \theta e^{-\theta x}$$

- Suppose that the distribution of population parameters is governed by a gamma distribution such that $\Theta \sim$ gamma(α, β)

$$f_\theta(\theta) = \frac{1}{\Gamma(\alpha)\beta^\alpha}\theta^{\alpha-1}e^{-\theta/\beta}$$

- The pdf of X is

$$\begin{aligned} f_X(x) &= \int f_{x|\theta}(x) f_\Theta(\theta) d\theta \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^\infty \theta^\alpha e^{-\theta(x+1/\beta)} d\theta = \frac{\alpha\beta}{(1+x\beta)^{\alpha+1}} \end{aligned}$$

- We recognize this as a Pareto distribution with parameters α and $\theta = 1/\beta$

Mixture Expectations

- For some mixtures such as the above example, we can compute the mixture distribution in closed-form
- However, it is often helpful to just consider the moments. For the mean function, using the *law of iterated expectations*, we have

$$\mathrm{E}\, X = \mathrm{E}_\Theta[\mathrm{E}(X|\Theta)]$$

- This is easily extended to the kth moment

$$\mathrm{E}\, X^k = \mathrm{E}_\Theta[\mathrm{E}(X^k|\Theta)]$$

Mixture Expectations Example

Example. Gamma Mixtures of Exponentials.

- Assume that $X|\theta \sim$ exponential with parameter $(\frac{1}{\theta})$. Thus, the mean is $\mathrm{E}(X|\theta) = 1/\theta$, the second raw moment is $\mathrm{E}(X^2|\theta) = 2/\theta^2$, and the variance is $\mathrm{Var}(X|\theta) = 1/\theta^2$
- For the parameter distribution, we have $\theta \sim$ gamma(α, β)
- One can check that

$$\mathrm{E}\, X = \frac{1}{\beta(\alpha - 1)}$$

and

$$\mathrm{Var}\, X = \frac{\alpha}{\beta^2(\alpha - 1)^2(\alpha - 2)}$$

- This is consistent with a Pareto distribution with parameters α and $\theta = 1/\beta$ (good practice to check)

Splicing

- Join (splice) together different probability density functions to form a pdf over the support of a random variable

$$f_X(x) = \begin{cases} \alpha_1 f_1(x) & c_0 < x < c_1 \\ \alpha_2 f_2(x) & c_1 < x < c_2 \\ \vdots & \vdots \\ \alpha_k f_k(x) & c_{k-1} < x < c_k \end{cases}$$

  $\alpha_1 + \alpha_2 \cdots + \alpha_k = 1$

  Each f_j is a pdf, so that $\int_{c_{j-1}}^{c_j} f_j(x)dx = 1$

  c_j's are typically known

- *Example: Life Contingencies*.
  - It is common to use an exponential distribution in the early ages, e.g., from $x = 5$ to $x = 40$. The exponential has a constant hazard rate and is well suited to model mortality from accidents
  - Beginning at age $x = 40$, one use another mortality law, e.g., Gompertz, that reflects mortality that increases with age x

Risk Retention Framework

- Now consider the following framework:
  - The policyholder or insured suffers a loss in the amount X
  - Under the insurance contract, the insurer is obligated to covered a portion of this amount
  - The insurer may have entered into a separate contract with a reinsurer that relieves the insurer of a portion of its obligations
- This section introduces standard mechanisms that insurers use to reduce, or mitigate, their risk, including deductibles and policy limits
- Further, we examine how the distribution of the insurers obligations depends on these mechanisms

Risk Retention Function

- Recall that X represents the amount of an insurable loss and use Y to represent the insurer's obligation
- There is a known function $g(\cdot)$ that maps the amount insured to the amount retained by the insurer, that is, $Y = g(X)$
- **Special Case 1. Deductible (d)**

$$g(x) = (x-d)_+ = \begin{cases} 0 & x \leq d \\ x-d & x > d. \end{cases}$$

  The notation "$(\cdot)_+$" means "take the positive part of." $Y = g(X)$ as the loss in excess of the deductible d

- **Special Case 2. Limit (u)**

$$g(x) = x \wedge u = \begin{cases} x & x \leq u \\ u & x > u. \end{cases}$$

  The notation "$\wedge$" means "take the minimum of." In this case, the insurance only pays up to a specified limit u. The random variable $Y = X \wedge u = \min(X, u)$ is the claim paid
- **Special Case 3. Coinsurance**. Define $Y = cX$. Typically, $0 < c < 1$, and so represents the proportion of claims retained by the insurer

Risk Retention Function II

- One handy way of combining the three special cases is through the expression

$$g(x) = \begin{cases} 0 & x \leq d \\ c(x-d) & d \leq x < u \\ c(u-d) & x \geq u. \end{cases}$$

- Think about these as parameters in a contract between a policyholder and an insurer and so represent "modifications" of the underlying contract
- Also interpret the risk retention function as the result from a reinsurance contract
  - For example, it is common in such a contract for an insurer to retain 50% of each risk and "cede" 50% to the reinsurer

Information Set for Deductibles

- Specify what type of information, sometimes known as the “information set,” that is available to the insurer
- **Special Case 4. Policyholder Deductible**. Define:

$$g_P(x) = \begin{cases} \text{undefined/not observed} & x \leq d \\ x - d & x > d \end{cases}$$

- The insurance only pays amounts in excess of the deductible d. If the loss is less than the deductible, then the insurer does not observe the loss. The random variable $Y^P = g(X)$ is the claim that an insurer observes
- We have placed a “P” subscript to remind ourselves that the retained loss is on what is sometimes known as a “per payment” basis
- In statistical terms, this retained loss is *truncated* in the sense that values of X below d are not observed
- To distinguish this from the other case where a zero is observed for losses $X < d$, the terminology **per loss** is used. Some sources use the notation $Y^L = (X - d)_+$ for the loss amount on a **per loss** basis

Distributions of Retained Risks - Deductible

- Consider two types of *ordinary deductible*:

1. Cost (amount of payment) per loss event

$$Y^L = (X-d)_+ \quad = \quad \begin{cases} 0 & X < d \\ X-d & X \geq d \end{cases} \quad \text{(a censored rv)}$$

2. Cost (amount of payment) per payment event

$$Y^P \quad = \quad \begin{cases} \mathit{undefined} & X < d \\ X-d & x \geq d \end{cases} \quad \text{(a truncated rv)}$$

Example. Exponential Distribution. Suppose that the loss X has distribution function $F(x) = 1 - \exp(-x/1000)$. Compute the distribution function and pdf for Y^L and Y^P with $d = 250$

Pareto Per Payment Deductible

- Assume a deductible $d = 1,000$

  Then, the claim amount on a per payment basis is

$$Y^P = \begin{cases} \text{undefined/not observed} & X < 1000 \\ 1000 & X \geq 1000 \end{cases}$$

- Identify the distribution of Y^P

Limited Expected Value

- Use a generic "u" for the upper limit. To compute the expected value of the limited loss variable $\min(X,u)$, we have

$$\mathrm{E}\ \min(X,u) = \int_0^u (1-F(x))\,dx = \int_0^u S(x)dx.$$

Pareto Policy Limit. Recall

$$1-F(x) = S(x) = \Pr(X > x) = \left(\frac{\theta}{x+\theta}\right)^{\alpha}$$

with mean $\mathrm{E}\ (X) = \frac{\theta}{\alpha-1}$. Thus, the limited expected value is

$$\begin{aligned}
\mathrm{E}\ \min(X,u) &= \theta^{\alpha}\int_0^u (x+\theta)^{-\alpha}dx = \theta^{\alpha}\left.\frac{(x+\theta)^{-\alpha+1}}{-\alpha+1}\right|_0^u \\
&= \theta^{\alpha}\left(\frac{\theta^{-\alpha+1}-(u+\theta)^{-\alpha+1}}{\alpha-1}\right) \\
&= \frac{\theta}{\alpha-1}\left\{1-\left(\frac{\theta}{u+\theta}\right)^{\alpha-1}\right\}.
\end{aligned}$$

Pareto Deductible

- The claim amount on a “per loss” basis is $Y^L = (X-d)_+$ for a deductible d
- To calculate $\mathrm{E}\,(X-d)_+$, we can use the relation, $X \wedge d + (X-d)_+ = X$
- For the Pareto distribution, recall $\mathrm{E}\,X = \frac{\theta}{\alpha-1}$ and

$$\mathrm{E}\ \min(X,d) = \frac{\theta}{\alpha-1}\left\{1-\left(\frac{\theta}{d+\theta}\right)^{\alpha-1}\right\}.$$

Thus,

$$\begin{aligned}\mathrm{E}\,(X-d)_+ &= \mathrm{E}\,X - \mathrm{E}\ \min(X,d) = \frac{\theta}{\alpha-1} - \frac{\theta}{\alpha-1}\left\{1-\left(\frac{\theta}{d+\theta}\right)^{\alpha-1}\right\}\\ &= \frac{\theta}{\alpha-1}\left\{\left(\frac{\theta}{d+\theta}\right)^{\alpha-1}\right\}.\end{aligned}$$

Mean Residual Life

- For the "per payment" random variable associated with the policyholder deductible case,

$$g_P(x) = \begin{cases} \text{undefined/not observed} & x \leq d \\ x - d & x > d \end{cases}$$

we can calculate the expectation as

$$e_X(d) \quad = \quad e(d) = \mathrm{E}\left(X - d | X > d\right)$$

- Here, $e_X(d)$ is known as the *mean residual life*
- We can write this as

$$\begin{aligned} e(d) &= \mathrm{E}\left(X - d | X > d\right) \\ &= \frac{\int_d^\infty (x-d) f(x) dx}{S(d)} = \frac{\mathrm{E}\left(X - d\right)_+}{S(d)} \end{aligned}$$

Thus,

$$e(d) \quad = \quad \frac{\int_d^\infty S(x) dx}{S(d)}$$

Example

Example. Exam M Fall 2005, Exercise 26. For an insurance:

1. Losses have density function

$$f_X(x) = \begin{cases} 0.02x & 0 < x < 10 \\ 0 & \text{elsewhere} \end{cases}$$

2. The insurance has an ordinary deductible of 4 per loss
3. Y^P is the claim payment per payment random variable

Calculate $\mathrm{E}\left[Y^P\right]$

Summary of Limited Loss Variables

Random Variable	**Expectation**
Excess loss random variable $Y = X - d$ if $X > d$ left *truncated*	$e_X(d) = \mathrm{E}\, Y = \mathrm{E}(X - d \mid X > d)$ mean excess loss function mean residual life function complete expectation of life $e_X^k(d) = \mathrm{E}\left[(X-d)^k \mid X > d\right]$
$(X-d)_+ = \begin{cases} 0 & X < d \\ X - d & X \geq d \end{cases}$ left-*censored* and shifted variable	$\mathrm{E}\ (X-d)_+ = e(d)S(d)$ $\mathrm{E}\ (X-d)_+^k = e^k(d)S(d)$
$\min(X,d) = X \wedge d = \begin{cases} X & X < d \\ d & X \geq d \end{cases}$ *limited loss variable* - right *censored*	$\mathrm{E}\,(X \wedge d)$ – limited expected value

Note that $(X-d)_+ + (X \wedge d) = X$. Thus, $\mathrm{E}\ (X-d)_+ + \mathrm{E}\,(X \wedge d) = \mathrm{E}\, X$
For nonnegative, continuous random variables,

$$\mathrm{E}\,(X \wedge d) = \int_0^d S(x)\, dx \quad \text{and} \quad \mathrm{E}(X-d)_+ = \int_d^\infty S(x)\, dx$$

Loss Elimination Ratio (LER)

- Consider an ordinary deductible, cost (amount of payment) per loss event

$$\begin{aligned} LER &= \frac{\mathrm{E}\,X - (\mathrm{E}\,X - \mathrm{E}\,(X \wedge d))}{\mathrm{E}\,X} = \frac{\mathrm{E}\,(X \wedge d)}{\mathrm{E}\,X} \\ &= \frac{\text{limited exp value}}{\text{exp value}} \end{aligned}$$

What fraction of losses have been eliminated by introducing the deductible?

Example. Losses have a lognormal distribution with $\mu = 6$ and $\sigma = 2$. There is a deductible of 2,000, and 10 losses are expected each year

Determine the loss elimination ratio