Probability and Stochastic Processes

Master in Actuarial Sciences

Alexandra Bugalho de Moura

2018/2019

Severity models (continuous models)

Severity models (continuous models)

Severity models

- Given that a claim occurs, the (individual) claim size X is typically referred to as claim severity.
- While typically this may be a continuous random variables, sometimes claim sizes can be considered discrete.
- When modelling claims severity, insurers are usually concerned with the tail of the distribution, specially for some contracts.
- In actuarial applications (to describe the severity distribution) we are mainly interested in distributions that have only positive support, i.e. $F_X(0) = 0$.
  Examples include the exponential; gamma; lognormal; Pareto, Burr and Inverse Gaussian.

Creating new distributions

Creating new distributions

There are many methods to generate new distributions; some of these methods allow us to give in-depth interpretation to the distributions.
The methods used can be sub-divided into:

- Addition of several random variables
  - For example, sums of (independent) Exponentials give a Gamma. This method will not be further explored.
- Transformation of random variables
  - Scalar multiplication.
  - Power operations.
  - Exponentiation (or logarithmic transformation).
- Mixing of distributions (we have already explored)
  - Frailty models
- Spliced distributions

Transformation of random variables

Multiplication by a constant or scalar transformation

- Let X be a continuous r.v. with pdf $f_X(x)$ and cdf $F_X(x)$
- Let $Y = aX$ with $a > 0$

$$F_Y(y) = F_X\left(\frac{y}{a}\right) \qquad \text{and} \qquad f_Y(y) = \frac{1}{a} f_X\left(\frac{y}{a}\right)$$

Multiplication by a constant or scalar transformation

- Insurance interpretation: if X denotes claims, then scalar transformation can be interpreted as applying an inflation factor across all levels of claims.
- A family of distributions that is closed under scalar multiplication (i.e. after scalar transformation, the new r.v. remains in the same family) is called a scale family of distributions.
- Some scale families are: Normal; Exponential; Pareto

Transformation of random variables

Power transformations

A power transformation involves raising the random variable by a power such as

$$Y = X^{1/\tau} \qquad \text{or} \qquad Y = X^{-1/\tau}, \qquad \tau > 0$$

- In the first case, we have a transformed X distribution;
- In the second case, we have an inverse transformed X distribution.
- In the special case where $Y = X^{-1}$, we have the inverse random variable.

Remarks

It is easy to show the following results, for $\tau > 0$:

- $Y = X^{1/\tau} \Longrightarrow F_Y(y) = F_X(y^\tau)$ and $f_Y(y) = \tau y^{\tau-1} f_X(y^\tau)$
- $Y = X^{-1/\tau} \Longrightarrow F_Y(y) = 1 - F_X(y^{-\tau})$ and $f_Y(y) = \tau y^{-\tau-1} f_X(y^{-\tau})$
- $Y = X^{-1} \Longrightarrow F_Y(y) = 1 - F_X(y^{-1})$ and $f_Y(y) = y^{-2} f_X(y^{-1})$

To retain θ as a scale parameter, the base distribution should be raised to a power before being multiplied by θ.

Transformation of random variables

Power transformations: Examples

Let $X \sim Exp(1)$

- $Y = \theta X^{-1}$ is the inverse exponential distribution, with $F_Y(y) = e^{-\theta/y}$, $y > 0$
- $Y = \theta X^{1/\tau}$ is the Weibull distribution, with $F_Y(y) = 1 - e^{-(y/\theta)^\tau}$, $y > 0$
- $Y = \theta X^{-1/\tau}$ is the inverse Weibull distribution, with $F_Y(y) = e^{-(\theta/y)^\tau}$, $y > 0$

Transformation of random variables

Exponentiation

Let X be a continuous r.v. with pdf $f_X(x)$ and cdf $F_X(x)$, and let $Y = \exp(X)$. Then

$$F_Y(y) = F_X(\ln y) \qquad \text{and} \qquad f_Y(y) = \frac{1}{y} f_X(\ln y)$$

Example

Let $X \sim N(\mu, \sigma)$. Derive the density of $Y = \exp(X)$ (lognormal distribution).

Transformation of random variables

General theory of transformation

Let X be a continuous r.v. with pdf $f_X(x)$ and cdf $F_X(x)$.
Let $Y = g(X)$ and assume tha g is a one-to-one transformation (i.e. invertible)

- If g is increasing, then
$$F_Y(y) = F_X(g^{-1}(y))$$
- If g is decreasing, then
$$F_Y(y) = 1 - F_X(g^{-1}(y))$$
- The density of Y is
$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right|$$
- If g is not monotone, we have to split g into monotone parts

Frailty Models

Frailty Models

A frailty model is a random effect model where the random effect (the frailty) has a multiplicative effect on the hazard It can be used to describe the influence of unobserved covariates in a proportional hazards model.

- Frailty models are particular type of mixture distributions
- Originally from the analysis of lifetime distributions in survival analysis
- May be viewed as a useful way to generate new distributions by mixing

Frailty Models

- Consider a frailty r.v. $\Lambda > 0$
- Define the conditional (given $\Lambda = \lambda$) hazard rate of X to be $h_{X|\Lambda=\lambda}(x) = \lambda a(x)$, where $a(x)$ is a specified known function of x
- The frailty is meant to quantify uncertainty associated with the hazard rate, acting in this case in a multiplicative manner

Frailty Models

Frailty Models

From $h_{X|\Lambda=\lambda}(x) = \lambda a(x)$, we have

$$S_{X|\Lambda=\lambda}(x) = e^{-\int_0^x h_{X|\Lambda=\lambda}(t)dt} = e^{-\int_0^x \lambda a(t)dt} = e^{-\lambda A(x)}$$

where $A(X) = \int_0^x a(t)dt$.

Frailty Models

The mixture distribution, i.e. the distribution of X, is defined by means of the mgf of Λ, $M_\Lambda(\cdot)$, assuming it exists:

$$S_X(x) = \int_0^\infty S_{X|\Lambda=\lambda}(x) f_\Lambda(\lambda)d\lambda = \int_0^\infty e^{-\lambda A(x)} f_\Lambda(\lambda)d\lambda = M_\Lambda(-A(x))$$

Frailty Models

Frailty Models: Example

- Let $X|\Lambda = \lambda \sim Weibull\left(\lambda^{-1/\tau}, \tau\right)$, i.e. conditional on $\Lambda = \lambda$, X has a Weibull distribution, with survival function

$$S_{X|\Lambda=\lambda}(x) = e^{-\lambda x^{\tau}}$$

that is, $A(x) = \int_0^x a(t)dt = x^{\tau}$ and $h_{X|\Lambda=\lambda}(x) = \lambda a(x)$.

- Let $\Lambda \sim Gamma(\alpha, \theta)$
- Then

$$S_X(x) = M_\Lambda(-A(x)) = M_\Lambda(-x^{\tau}) = (1 + \theta x^{\tau})^{-\alpha}$$

- Hence $X \sim Burr\left(\alpha, \theta^{-1/\tau}, \tau\right)$

Spliced distributions

Spliced distributions

- A spliced distribution is one whose form of distribution is different in different portions of the domain of the random variable.
- An interpretation in insurance claims is that the distributions vary by size of claims.
- To illustrate, consider a two-spliced distribution:

$$f_X(x) = \begin{cases} p\, f_1(x), & \text{for} \quad 0 < x < c \\ (1-p)\, f_2(x), & \text{for} \quad c < x < +\infty \end{cases}$$

  where f_1 and f_2 are both legitimate density functions on the corresponding intervals.
- This concept can be extended to a k-component spliced distributions.

Example

Create a two component spliced model using an exponential with parameter θ from 0 to c and a $Pareto(\alpha, \gamma)$ from c to ∞.

Normal distribution

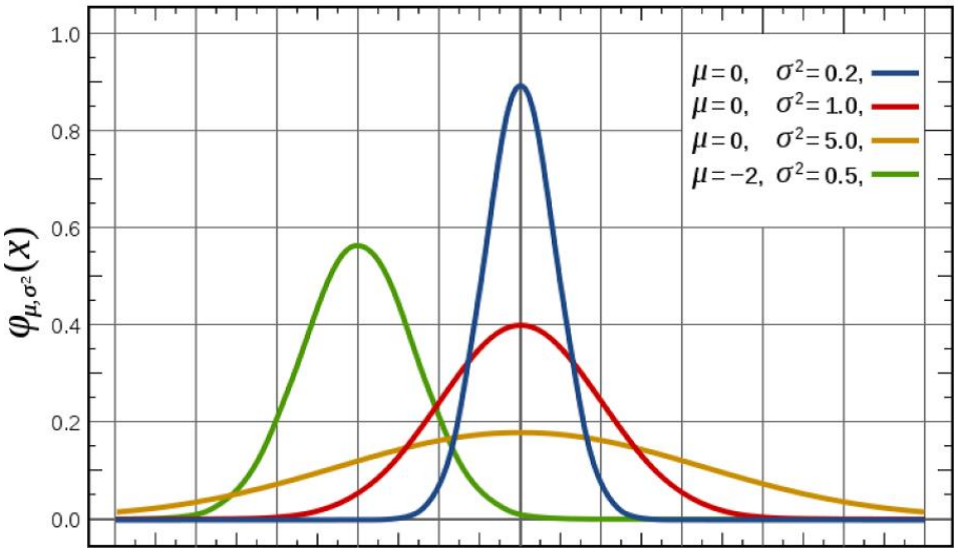
$X \sim N(\mu, \sigma), \quad -\infty < \mu < +\infty, \quad \sigma > 0$

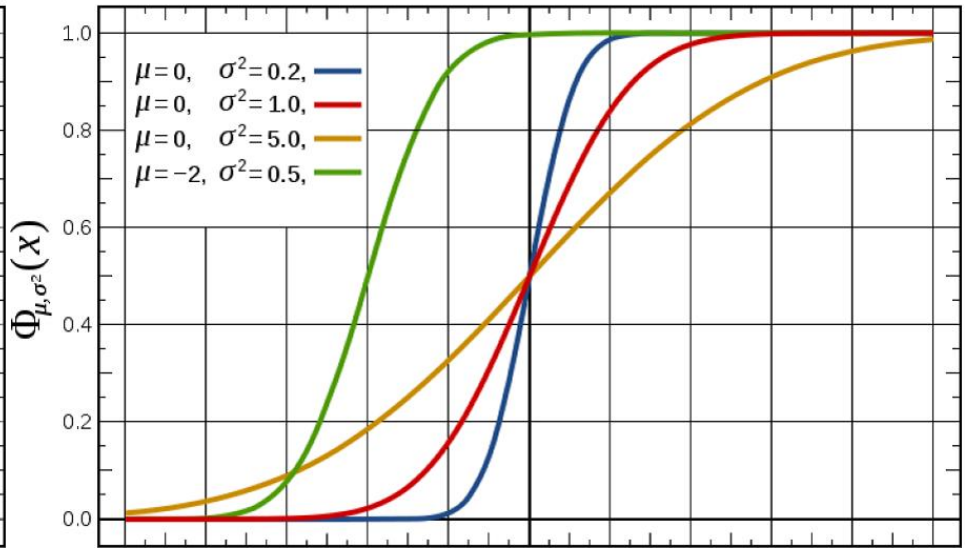
- μ is a location parameter and σ is a shape parameter (not a scale parameter unless $\mu = 0$)

$f_X(x)$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty$
$F_X(x)$	$\int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt, \quad -\infty < x < +\infty$
$M_X(t)$	$e^{t\mu + \frac{t^2\sigma^2}{2}}$
$E[X]$	μ
$V[X]$	σ^2
μ_k	$\begin{cases} 0, \text{ if } k \text{ is odd} \\ \frac{k!}{(k/2)!}\frac{\sigma^k}{2^{k/2}} \text{ if } k \text{ is even} \end{cases}$

Normal distribution

$X \sim N(\mu, \sigma), \quad -\infty < \mu < +\infty, \quad \sigma > 0$

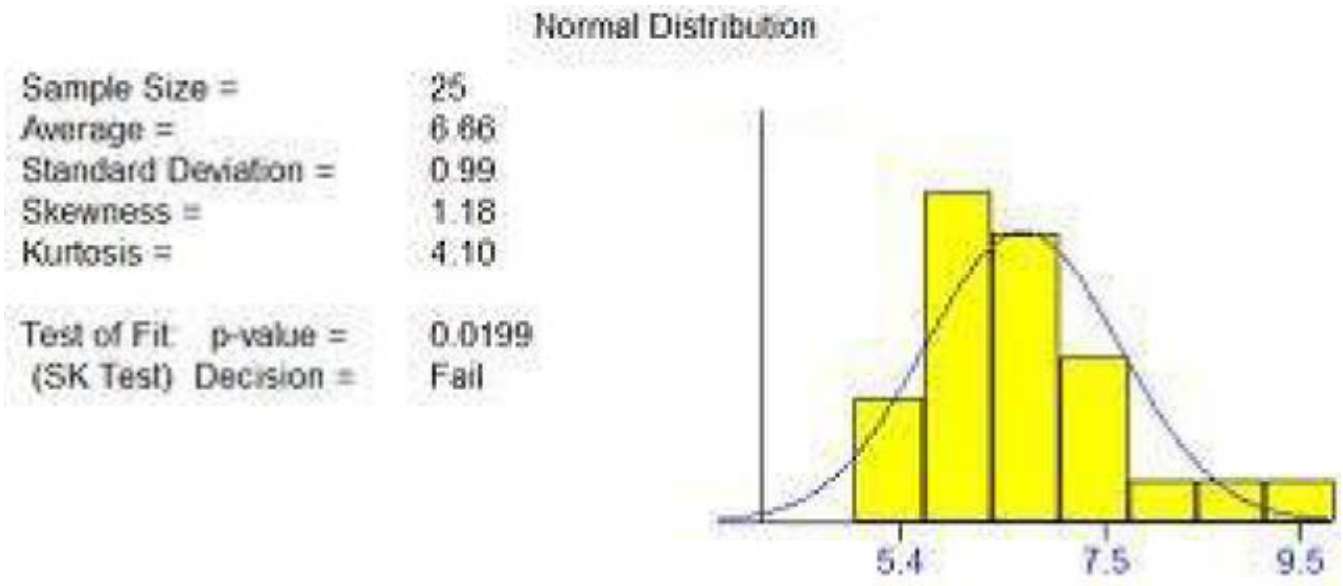




Normal distribution

$X \sim N(\mu, \sigma), \quad -\infty < \mu < +\infty, \quad \sigma > 0$

- Standard Normal: when $\mu = 0$ and $\sigma = 1$; $X \sim N(\mu, \sigma) \Longrightarrow Z = \dfrac{X - \mu}{\sigma} \sim N(0, 1)$
- Sum of Normal random variables is Normal
- $X \sim N(\mu, \sigma) \Longrightarrow cX \sim N(c\mu, c\sigma)$
- Careful about using the Normal



Gamma distribution

$X \sim Gamma(\alpha, \theta), \quad \alpha, \theta > 0$

- θ is a scale parameter and α is a shape parameter

$f_X(x)$	$\frac{1}{\Gamma(\alpha)} \frac{x^{\alpha-1}}{\theta^\alpha} e^{-x/\theta}, \quad x > 0$
$F_X(x)$	$\int_0^x \frac{1}{\Gamma(\alpha)} \frac{t^{\alpha-1}}{\theta^\alpha} e^{-t/\theta} dt, \quad x > 0$
$M_X(t)$	$(1 - t\theta)^{-\alpha}, \ t < 1/\theta$
$E[X^k]$	$\frac{\theta^k \Gamma(\alpha + k)}{\Gamma(\alpha)}$
$V[X]$	$\alpha\theta^2$
Mode$[X]$	$\theta(\alpha - 1), \quad \alpha > 1$

Gamma distribution

$X \sim Gamma(\alpha, \theta), \quad \alpha, \theta > 0$

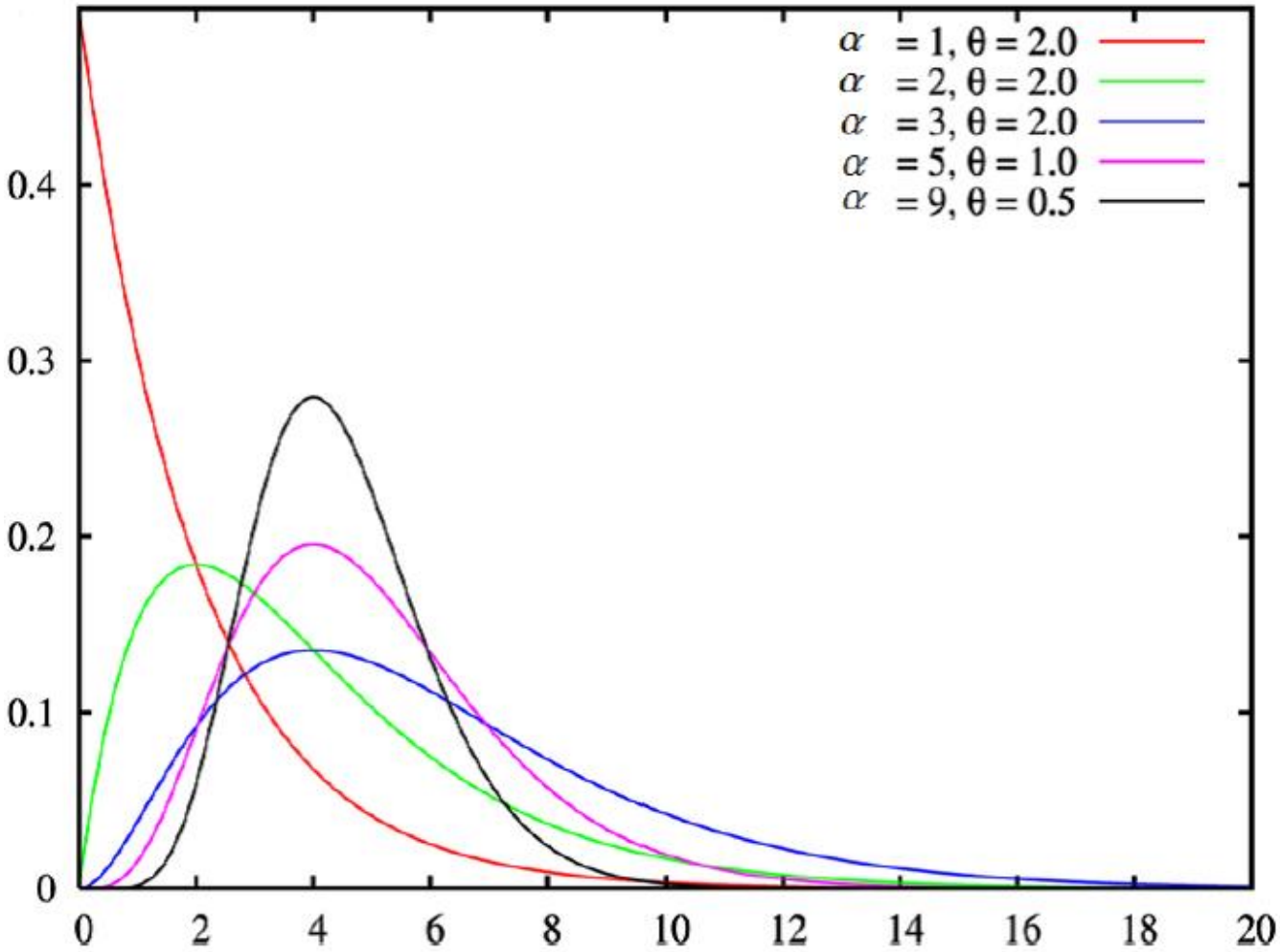
- Use it if the tail of the distribution is considered light (the mgf exists)
- Applicable, e.g., to damage to automobiles
- $X \sim Gamma(\alpha, \theta) \Longrightarrow Y = \lambda X \sim Gamma(\alpha, \lambda\theta)$
- $X_i \sim Gamma(\alpha_i, \theta)$ indep. $\Longrightarrow S = \sum_{i=1}^{n} X_i \sim Gamma\left(\sum_{i=1}^{n} \alpha_i, \theta\right)$

Special cases

- **Exponential:** $\alpha = 1 \Longrightarrow X \sim Exp(\theta)$
- **Chi-square** $\alpha = n/2$ and $\theta = 2 \Longrightarrow X \sim \chi^2_{(n)}$
- $X \sim Gamma(n/2, \theta) \Longrightarrow Y = 2X/\theta \sim Gamma(n/2, 2) = \chi^2_{(n)}$

Gamma distribution

$X \sim Gamma(\alpha, \theta), \quad \alpha, \theta > 0$



Lognormal distribution

$X \sim Lognormal(\mu, \sigma), \quad -\infty < \mu < +\infty, \quad \sigma > 0$

$f_X(x)$	$\frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad x > 0$
$F_X(x)$	$\Phi\left(\frac{\ln x - \mu}{\sigma}\right), \quad x > 0$
$M_X(t)$	Does not exist for $t > 0$
$E[X^k]$	$e^{k\mu + \frac{1}{2}k^2\sigma^2}$
$V[X]$	$e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right)$
Mode$[X]$	$e^{\mu - \sigma^2}$
Med$[X]$	e^{μ}

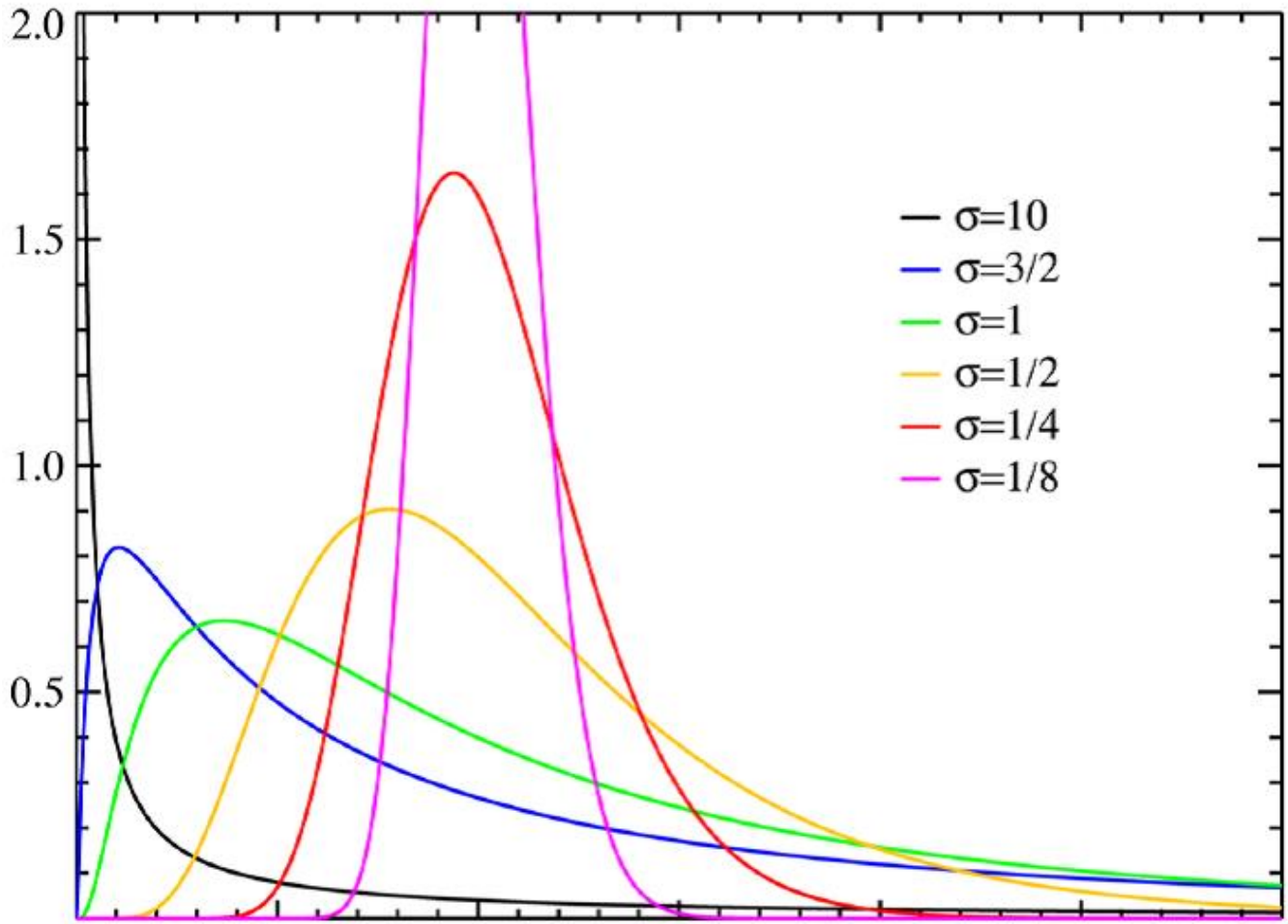
Lognormal distribution

$X \sim Lognormal(\mu, \sigma), \quad -\infty < \mu < +\infty, \quad \sigma > 0$

- Havier tailed than the Gamma (although all moments exist, the mgf does not exist)
- Applicable for example to fire insurance
- $Y \sim Normal(\mu, \sigma) \Longrightarrow X = exp(Y) \sim Lognormal(\mu, \sigma)$
- $X \sim Lognormal(\mu, \sigma) \Longrightarrow Y = \ln X \sim N(\mu, \sigma)$
- The product of independent Lognormals is also Lognormal
- A lognormal distribution is not uniquely determined by its moments $E[X^k]$ for $k \geqslant 1$, that is, there exists some other distribution with the same moments for all k.

Lognormal distribution

$X \sim LogNormal(\mu, \sigma), \quad -\infty < \mu < +\infty, \quad \sigma > 0$



Pareto distribution

$X \sim Pareto(\alpha, \theta), \quad \alpha, \theta > 0$

- θ is a scale parameter and α is a shape parameter
- Used for heavy tailed business, such as liability insurance
- It is a continuous mixture of exponentials with Gamma mixing weights
- $X \sim Pareto(\alpha, \theta)$, and $c > 0 \Longrightarrow cX \sim Pareto(\alpha, c\theta)$
- We can calculate $e_X(d)$ as the expected value of a pareto with parameters α and $d + \theta$:

$$X \sim Pareto(\alpha, \theta) \Longrightarrow Y^P \sim Pareto(\alpha, d + \theta) \quad \left(\text{remind that } S_d(x) = \frac{S_X(x+d)}{S_X(d)} \right)$$

- $X \sim Pareto(\alpha, \theta)$ and $Y = X + \theta$, then

$$f_Y(y) = \frac{\alpha \theta^\alpha}{y^{\alpha+1}}, \quad y > \theta$$

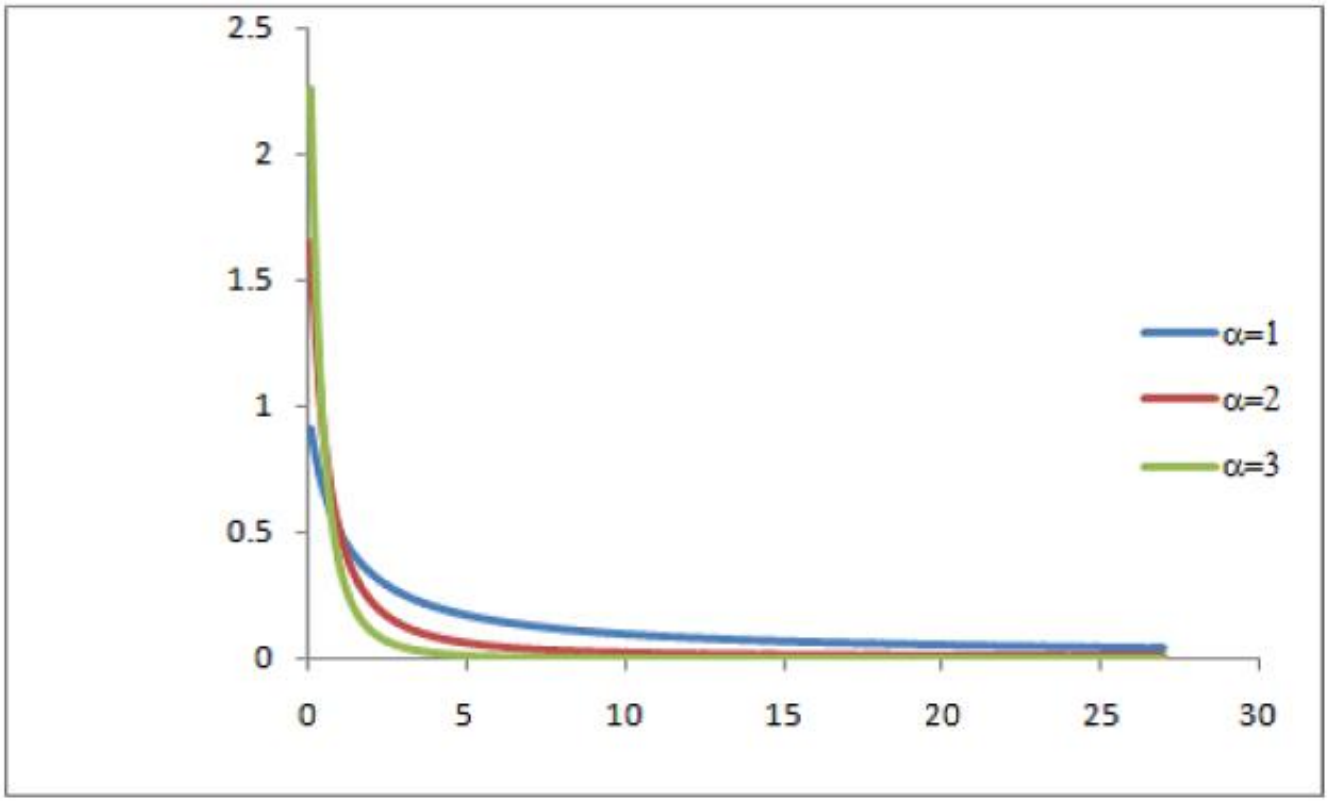
Y is called, in Loss Models book, the **single-parameter Pareto**, assuming the value of θ is known

$X \sim Pareto(\alpha, \theta), \quad \alpha, \theta > 0$

$f_X(x)$	$\frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}, \quad x > 0$
$F_X(x)$	$1 - \left(\frac{\theta}{x+\theta}\right)^\alpha, \quad x > 0$
$M_X(t)$	Does not exist
$E[X^k]$	$\theta^k k! \frac{\Gamma(\alpha-k)}{\Gamma(\alpha)}, \quad k < \alpha$
$V[X]$	$\frac{\alpha\theta^2}{(\alpha-2)(\alpha-1)^2}, \quad \alpha > 2$
$h_X(x)$	$\frac{\alpha}{x+\theta}, \quad x > 0$
$e_X(x)$	$\frac{x+\theta}{\alpha-1}, \quad x > 1$
$E[X \wedge x]$	$\begin{cases} \frac{\theta}{\alpha-1}\left(1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha-1}\right), & \text{if } \alpha \neq 1 \\ -\theta \ln\left(\frac{\theta}{x+\theta}\right) \text{ if } \alpha > 1 \end{cases}$

Pareto distribution

$X \sim Pareto(\alpha, \theta), \quad \alpha, \theta > 0$



Burr distribution

$X \sim Burr(\alpha, \theta, \gamma), \quad \alpha, \theta, \gamma > 0$

- The Burr Type XII distribution or simply the Burr distribution is a continuous probability distribution for a non-negative random variable.
- It is also known as the Singh-Maddala distribution and is one of a number of different distributions sometimes called the "generalized log-logistic distribution".
- In insurance it is used for heavy-tailed business, such as liability insurance.

Burr distribution

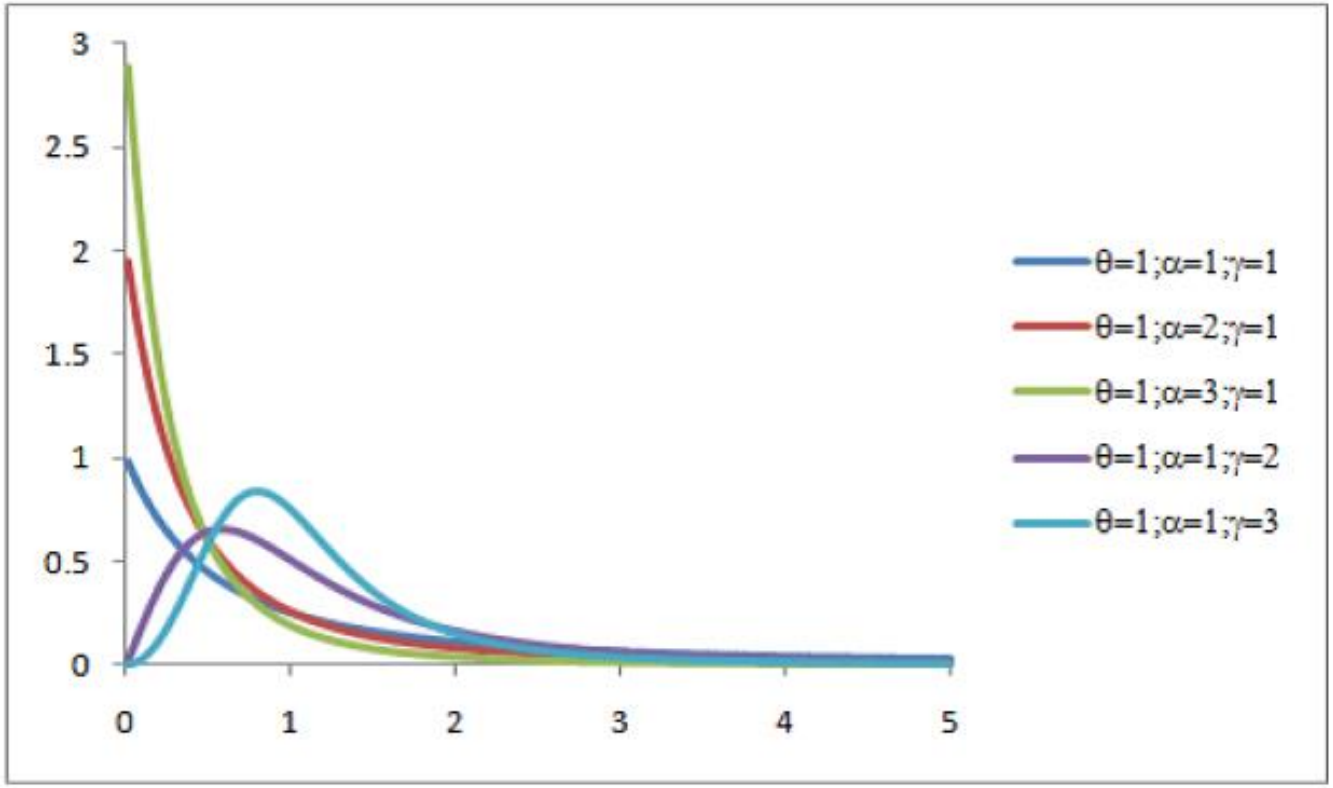
$X \sim Burr(\alpha, \theta, \gamma), \quad \alpha, \theta, \gamma > 0$

- When $\gamma = 1$ we obtain a Pareto

$f_X(x)$	$\frac{\alpha\gamma(x/\theta)^\gamma}{x(1+(x/\theta)^\gamma)^{\alpha+1}}, \quad x > 0$
$F_X(x)$	$1 - \left(\frac{1}{1+(x/\theta)^\gamma}\right)^\alpha, \quad x > 0$
$M_X(t)$	Does not exist
$E[X^k]$	$\theta^k \frac{\Gamma(1+k/\gamma)\Gamma(\alpha-k/\gamma)}{\Gamma(\alpha)}, \quad -\gamma < k < \alpha\gamma$
Mode$[X]$	$\begin{cases} \theta\left(\frac{\gamma-1}{\alpha\gamma+1}\right)^{1/\gamma} & \text{if } \gamma > 1 \\ 0 & \text{if } \gamma \leqslant 1 \end{cases}$

Burr distribution

$X \sim Burr(\alpha, \theta, \gamma), \quad \alpha, \theta, \gamma > 0$



Beta distribution

$X \sim Beta(\alpha, \beta)$

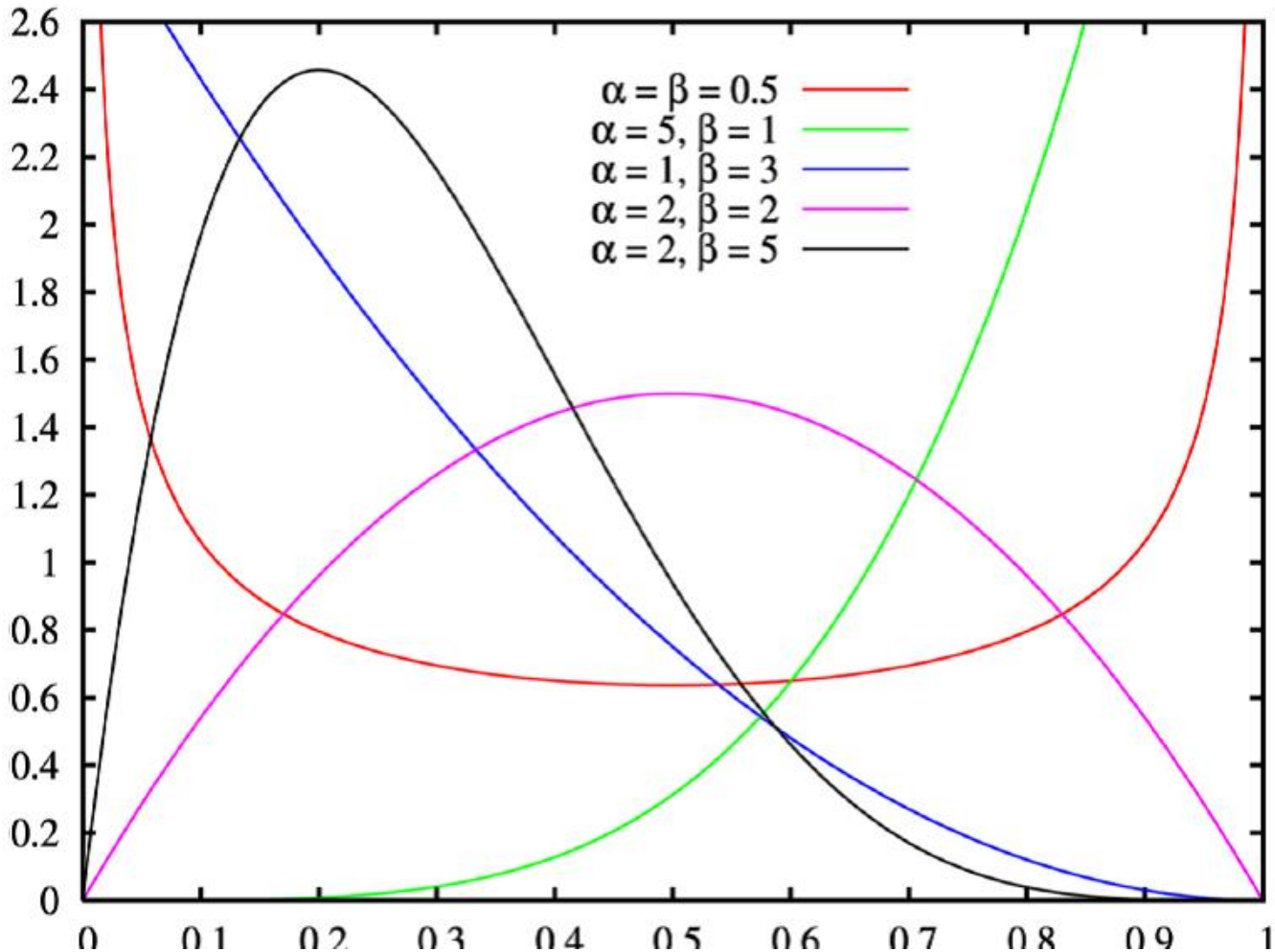
$f_X(x)$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha}(1-x)^{\beta-1}\frac{1}{x}, \quad 0 < x < 1$
$E[X^k]$	$\frac{\Gamma(\alpha+\beta)\Gamma(\alpha+k)}{\Gamma(\alpha)\Gamma(\alpha+\beta+k)}, \quad k > -\alpha$

$X \sim Beta(\alpha, \beta)$

Derive the density function of $Y = \theta X$ to obtain a $Beta(\alpha, \beta, \theta)$

Beta distribution

$X \sim Beta(\alpha, \beta)$



Extreme Value Theory

Extreme Value Theory (Loss Models)

- Sometimes actuaries are only interested in the distribution of large losses, e.g., for the per-claim reinsurance arrangements.
- Extreme Value Theory (EVT): supports the choice of the models to be used in these situations.
- EVT is concerned with two types of loss:
  - The largest loss over a period of time (less relevant for actuarial applications; more important for operational risk assessment)
  - Distribution of losses in excess of a threshold (directly relevant to actuarial work, e.g. reinsurance).
- Key results in EVT
  - The limiting distribution of the largest observation must be one of a very small number of distributions
  - The limiting distribution of the excess over a threshold mus be one of a small number of distributions
  - The shape of the distribution from which the sample is drawn determines which one of the distributions is appropriate
- This convenient theory allows us to rationally extrapolate to loss amounts that are well in excess of any historic loss and thus gives an idea of the magnitude of probabilities of large losses, even when those losses have never before occurred.

Extreme Value Distributions

Extreme Value Distributions

There are three related distributions in the familly known as extreme value distributions:

- Gumbel
- Fréchet
- Weibull

Extreme Value Distributions

Gumbel distribution

The standardized Gumbel distribution has cdf

$$F_X(x) = G_0(x) = e^{-e^{-x}}, \quad -\infty < x < +\infty$$

With location parameter μ and scale parameter θ, it becomes

$$F_X(x) = G_{0,\mu,\theta}(x) = e^{-e^{-\left(\frac{x-\mu}{\theta}\right)}}, \quad -\infty < x < +\infty, \quad \theta > 0$$

Extreme Value Distributions

Fréchet distribution

The standardized Fréchet distribution has cdf

$$F_X(x) = G_{1,\alpha}(x) = e^{-x^{-\alpha}}, \quad x \geqslant 0, \quad \alpha > 0$$

with shape parameter α.
With location parameter μ and scale parameter θ, it becomes

$$F_X(x) = G_{1,\alpha,\mu,\theta}(x) = e^{-\left(\frac{x-\mu}{\theta}\right)^{-\alpha}}, \quad x \geqslant \mu, \quad \alpha, \theta > 0$$

Remarks

- This distribution has support only for values of x greater than the location parameter μ
- When μ is set to zero it is a two-parameter distribution: $G_{1,\alpha,0,\theta}(x)$
- The two-parameter Fréchet distribution is the inverse Weibull distribution

Extreme Value Distributions

Weibull distribution

The standardized Weibull distribution has cdf

$$F_X(x) = G_{2,\alpha}(x) = e^{-(-x)^{-\alpha}}, \quad x \leqslant 0, \quad \alpha < 0$$

with shape parameter α.
With location parameter μ and scale parameter θ, it becomes

$$F_X(x) = G_{2,\alpha,\mu,\theta}(x) = e^{-\left(-\frac{x-\mu}{\theta}\right)^{-\alpha}}, \quad x \leqslant \mu, \quad \alpha < 0$$

Remarks

- This Weibull distribution is not the one we have seen before
- It has support only for values of x smaller than the location parameter μ
- It is often associated with the distribution of the minimum values of distributions and with distributions that have a finite righthand endpoint of the support of the distribution.
- Because insurance losses rarely have these characteristics, this model is not discussed further

Extreme Value Distributions

Generalized extreme value distribution

Family of distributions incorporating the Gumbel, the Fréchet and the Weibull distributions as special cases.

$$F_X(x) = G_\gamma(x) = \exp\left[-(1+\gamma x)^{-1/\gamma}\right]$$

Remarks

- $\lim_{\gamma \to 0} G_\gamma = \exp\left(-e^{-x}\right) = G_0(x)$ (Gumbel standardized distribution)
- When $\gamma > 0$, the cdf of $G_\gamma(x)$ takes the form of a Fréchet distribution
- When $\gamma < 0$, the cdf of $G_\gamma(x)$ takes the form of a Weibull distribution

Distribution of the maximum

Distribution of the maximum: from a fixed number of losses n

- Consider a set of n (fixed) observations of iid non-negative r.v. with distribution $F_X(x)$.
- Let M_n denote the maximum value of the n observations, with corresponding cdf and pdf $F_n(x)$ and $f_n(x)$. Then

$$F_n(x) = [F_X(x)]^n$$

Remarks

- In this case the cdf of the maximum is a simple function of the common distribution of the original random variables.
- The limiting distribution of the maximum in this case is degenerate: as $n \to \infty$, $F_n(x)$ approaches either 0 or 1, depending on whether $F_X(x) < 1$ or $F_x(x) = 1$.
- To avoid the effect of degeneracy in the limit, the study of the behaviour of the maximum requires appropriate normalization
- For distributions with no upper limit of support, this maximum continues to increase without limit as $n \to \infty$
- For distributions with a right-hand endpoint, the maximum approaches that right-hand endpoint as $n \to \infty$

Distribution of the maximum

Mean of the maximum: from a fixed number of losses n

For nonnegative random variables, the mean (if it exists) of the maximum is

$$E[M_n] = \int_0^\infty S_n(x)\,dx = \int_0^\infty \left(1 - [F_X(x)]^n\right)\,dx$$

Second raw moment of the maximum: from a fixed number of losses n

$$E[M_n^2] = 2\int_0^\infty xS_n(x)\,dx = 2\int_0^\infty x\left(1 - [F_X(x)]^n\right)\,dx$$

Distribution of the maximum

Example

- Suppose that we have carried out studies of the largest losses over many months
- Determine the distribution of the annual maximum assuming that the cdf of the monthly maximum follows a Gumbel distribution $G_{0,\mu,\theta}(x)$
- Determine the distribution of the annual maximum supposing now that the monthly maximum follows a Fréchet distribution $G_{1,\alpha,\mu,\theta}(x)$

Distribution of the maximum

Distribution of the maximum: from a random number N of losses

- In most cases, the number of losses in a period will fluctuate and thus it is a random variable
- Denote the random number of losses by N and its pgf by $P_N(z)$

Using the law of total probability:

$$\begin{aligned} F_N(x) &= P(M_N \leqslant x) = \sum_{n=0}^{\infty} P(M_N \leqslant x | N = n)\, P(N = n) = \sum_{n=0}^{\infty} [F_X(x)]^n\, P(N = n) \\ &= P_N(F_X(x)) \end{aligned}$$

Remarks

- If we can specify the distribution of the frequency and severity of losses, we can easily obtain the exact distribution of the maximum loss
- This distribution has a jump at $x = 0$ with value $P_N(F_X(0))$, the probability of no loss cost (either no loss event occurs, or all loss events have no cost)
- If $F_X(0) = 0$ then $P_N(F_X(0)) = P_N(0) = P(N = 0)$: the jump at zero is the probability that no loss occurs

Distribution of the maximum

Example

- Consider that the number of claims per year follows a **Poisson** process with a rate of λ losses per year.
- What is the general cdf of the maximum loss per year?
- What is the general cdf of the maximum loss in a period of k years?
- If the individual claim severity is exponentially distributed, what is the cdf of the maximum loss? (Gumbel)
- If the individual losses are Pareto, what is the cdf of the maximum loss? (Fréchet)

The Gumbel and Fréchet distributions are distributions of extreme statistics, in this case maxima. The Weibull plays the corresponding role for minima.

Example

- Suppose that the number of losses in one year follows a **negative binomial** distribution with parameters r and β.
- What is the general cdf of the maximum loss?
- Supposing that losses are exponentially distributed, what is the cdf of the maxium loss?
- Supposing that losses are Pareto distributed, what is the cdf of the maximum loss?

Stability of the maximum of the extreme value distribution

Stability of the maximum of the extreme value distribution

The extreme value distributions have the property of the **stability of the maximum** or **max-stability**, that is very useful in extreme value theory:

- If losses follow an extreme value distribution, then the maximum of n (fixed) observations has the same extreme value distribution, after a location or scale normalization.

Stability of the maximum of the extreme value distribution

Gumbel distribution

The distribution of the maximum of n observations from the standardized Gumbel distribution has itself a Gumbel distribution, after a shift of location:

$$[G_0(x + \ln n)]^n = G_0(x) \qquad \text{and equivalently} \qquad [G_0(x)]^n = G_0(x - \ln n)$$

Including location and scale parameters yields:

$$\left[G_{0,\mu,\theta}(x)\right]^n = G_{0,\mu^*,\theta}(x), \qquad \text{with } \mu^* = \mu + \theta \ln n$$

Stability of the maximum of the extreme value distribution

Fréchet distribution

The distribution of the maximum of n observations from the standardized Fréchet distribution has itself a Fréchet distribution, after a scale change:

$$\left[G_{1,\alpha}(n^{1/\alpha}x)\right]^n = G_{1,\alpha}(x) \qquad \text{and equivalently} \qquad [G_{1,\alpha}(x)]^n = G_{1,\alpha}\left(\frac{x}{n^{1/\alpha}}\right)$$

Including location and scale parameters yields:

$$\left[G_{1,\alpha,\mu,\theta}(x)\right]^n = G_{1,\alpha,\mu^*,\theta}\left(\frac{x-\mu}{\theta\, n^{1/\alpha}}\right) = G_{1,\alpha,\mu,\theta^*}(x), \qquad \text{with } \theta^* = \theta\, n^{1/\alpha}$$

Fisher-Tippett Theorem

Distribution of the maximum of n fixed observations, as $n \to \infty$

- The extreme value distributions are the limiting distributions (as $n \to \infty$) of extreme statistics for any distribution.
- Namely, they approximate distributions of the maximum (of n fixed observations) for (almost) any distribution
- As $n \to \infty$, the distribution of the maximum of n fixed observations is degenerate
- Hence, to understand the shape of the distribution for large values of n we need to normalize the r.v. representing the maximum
- We require linear transformations, *i.e.* sequences a_n and b_n, s.t. $\dfrac{M_n - b_n}{a_n}$ has (non-degenerate) limiting distribution:
  (we control for the growth of M_n)

$$\lim_{n\to\infty} F_n\left(a_n x + b_n\right) = \lim_{n\to\infty} P\left(M_n \leqslant a_n x + b_n\right) = G(x)$$

  where $G(x)$ is a nondegenerate distribution
- If such transformation exists, the Fisher-Tippett theorem provides a powerful result

Fisher-Tippett Theorem

Fisher-Tippett Theorem

If $[F_X(a_m x + b_n)]^n$ has a nondegenerate limiting distribution as $n \to \infty$ for some constants a_n and b_n that depend on n, then

$$\lim_{n\to\infty} [F_X(a_n x + b_n)]^n = G(x), \quad \forall x$$

where G is an extreme value distribution which is one of G_0, $G_{1,\alpha}$ or G_2, α for some location and scale parameters.

Remarks

- If we are interested in understanding how large losses behave, we only need to look at three (two, since Weibull has an upper limit) choices for a model for the extreme right-hand tail
- The Fisher-Tippett theorem requires normalization using appropriate norming constants a_n and b_n that dependen on n. For specific distributions, these norming constants can be identified.
- The FisherTippett theorem is a limiting result that can be applied to any distribution $F_X(x)$. Because of this, it can be used as a general approximation to the true distribution of a maximum without having to completely specify the form of the underlying distribution $F_X(x)$.

Fisher-Tippett Theorem

Example: maximum of exponentials

Let $X \sim Exp(1)$. Considering the norming constants $a_n = 1$ and $b_n = \ln n$, the distribution of the maximum is

$$\begin{aligned} P(M_n \leqslant a_n x + b_n) &= [P(X \leqslant a_n x + b_n)]^n = [P(X \leqslant x + \ln n)]^n \\ &= \left[1 - e^{-x-\ln n}\right]^n = \left[1 - \frac{e^{-x}}{n}\right]^n \longrightarrow e^{-e^{-x}} \quad \text{as } n \to \infty \end{aligned}$$

Having chosen the right norming constants, we see that the limiting distribution of the maximum of exponential random variables is the Gumbel distribution.

Fisher-Tippett Theorem

Example: maximum of Paretos

Let $X \sim Pareto(\alpha, \theta)$, i.e. with survival function $S(x) = \left(1 + \frac{x}{\theta}\right)^{-\alpha}$, $x \geqslant 0$ and $\alpha, \theta > 0$.

Considering the norming constants $a_n = \frac{\theta n^{1/\alpha}}{\alpha}$ and $b_n = \theta n^{1/\alpha} - \theta$, the distribution of the maximum is

$$\begin{aligned} P(M_n \leqslant a_n x + b_n) &= [P(X \leqslant a_n x + b_n)]^n = \left[P\left(X \leqslant \frac{\theta n^{1/\alpha}}{\alpha} x + \theta n^{1/\alpha} - \theta\right)\right]^n \\ &= \left[1 - \left(1 + \frac{\frac{\theta n^{1/\alpha}}{\alpha} x + \theta n^{1/\alpha} - \theta}{\theta}\right)^{-\alpha}\right]^n \\ &= \left[1 - \frac{1}{n}\left(1 + \frac{x}{\alpha}\right)^{-\alpha}\right]^n \longrightarrow e^{-\left(1+\frac{x}{\alpha}\right)^{-\alpha}} \quad \text{as } n \to \infty \end{aligned}$$

The limit as $n \to \infty$ of the maximum of Pareto random variables has a Frchet distribution with $\mu = -\alpha$ and $\theta = \alpha$.

Generalized Pareto Distributions

Generalized Pareto (GP) Distributions

- The GP distributions here introduzed are closely related to EV distributions
- They are mainly used when studying excesses over a threshold
- The GP distribution $W(x)$ is given by

$$W(x) = 1 + \ln G(x)$$

where

  - $G(x)$ is an EV distribution
  - $0 \leqslant W(x) \leqslant 1$, thus we require that $G(x) \geqslant e^{-1}$

- In the same way the EV distributions $G(x)$ may be of three types, there are three related distributions in the family of GP distributions:
  - Exponential
  - Pareto
  - Beta

Remark (Loss Models)

- The GP distribution referred here differs from the distribution with the same name given in Appendix A

Generalized Pareto Distributions

Exponential distribution

The standardized exponential distribution has cdf

$$F(x) = W_0(x) = 1 + \ln G_0(x) = 1 - e^{-x}, \qquad x > 0$$

With location and scale parameters μ and θ included, it has cdf

$$F(x) = W_{0,\mu,\theta}(x) = 1 + \ln G_{0,\mu\theta}(x) = 1 - e^{-\frac{x-\mu}{\theta}}, \quad x > \mu$$

Remarks

- The exponential distribution has support only for values of x greater than μ
- Commonly $\mu = 0$, making it a one-parameter distribution:

$$F(X) = W_{0,\theta}(x) = 1 - e^{-x/\theta}, \quad x > 0$$

Generalized Pareto Distributions

Pareto distribution

The standardized Pareto distribution has cdf

$$F(x) = W_{1,\alpha}(x) = 1 - \ln G_{1,\alpha}(x) = 1 - x^{-\alpha}, \qquad x \geqslant 1, \quad \alpha > 0$$

With location and scale parameters μ and θ included, it has cdf

$$F(x) = W_{1,\alpha,\mu,\theta}(x) = 1 + \ln G_{1,\alpha,\mu,\theta}(x) = 1 - \left(\frac{x-\mu}{\theta}\right)^{-\alpha}, \quad x \geqslant \mu + \theta, \quad \alpha, \theta > 0$$

Remarks

- The Pareto distribution has support only for values of x greater than $\mu + \theta$
- Commonly $\mu = -\theta$, making it a two-parameter distribution:

$$F(x) = W_{1,\alpha,\theta} = 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha}, \quad x \geqslant 0, \quad \alpha, \theta > 0$$

- The case $\mu = 0$ is denoted single-parameter Pareto distribution in the appendix

Generalized Pareto Distributions

Beta distribution

The standardized Beta distribution has cdf

$$F(x) = W_{2,\alpha}(x) = 1 + \ln G_{2,\alpha}(x) = 1 - (-x)^{-\alpha}, \qquad -1 \leqslant x \leqslant 0, \quad \alpha < 0$$

With location and scale parameters μ and θ included, it has cdf

$$F(x) = 1 - \left(-\frac{x-\mu}{\theta}\right)^{-\alpha}, \quad \mu - \theta \leqslant x \leqslant \mu, \quad \alpha < 0, \quad \theta > 0$$

Remarks

- The Beta distribution has support only for values of $x \in [\mu - \theta, \mu]$
- It is a (shifted) subclass of the usual beta distribution on the interval $(0, 1)$, with an additional location parameter, and where shape parameters are positive

Generalized Pareto Distributions

Generalized Pareto Distributions

- The GP distribution is the family of distributions incorporating, in a single expression, the preceding three distributions as special cases.
- The general expression for the cdf of the GP distribution is

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

- $W_{0,\theta}(x) = \lim_{\gamma \to 0} W_{\gamma,\theta}(x) = \lim_{\gamma \to 0} 1 - \left(1 + \gamma \frac{x}{\theta}\right)^{-1/\gamma} = 1 - e^{-x/\theta}$: exponential distribution.
- When $\gamma > 0$, the cdf $W_{\gamma,\theta}(x)$ is the Pareto distribution

Stability of excesses of the generalized Pareto

Stability of excesses of the generalized Pareto distributions

The exponential, Pareto and beta distributions have the property of **stability of excesses**

- The conditional distribution of the excess over a threshold of a generalized Pareto is of the same form as the underlying distribution
- Let $Y = X - d | X > d$ denote the excess loss r.v.
- If $X \sim W_{\gamma,\theta}(x)$ then

$$\begin{aligned} F_Y(y) &= P(Y \leqslant y) = P(X \leqslant d + y | X > d) = 1 - \frac{S_X(d+y)}{S_X(d)} \\ &= 1 - \left(\frac{1 + \gamma \left(\frac{d+y}{\theta} \right)}{1 + \gamma \left(\frac{d}{\theta} \right)} \right)^{-1/\gamma} = 1 - \left(1 + \gamma \frac{y}{\theta + \gamma\, d} \right)^{-1/\gamma} \\ &= W_{\gamma, \theta + \gamma d}(y), \quad y > 0 \end{aligned}$$

Stability of excesses of the generalized Pareto

Stability of excesses of the **exponential**

If $X \sim Exp(\theta)$, $F_X(x) = W_{0,\theta}(x) = 1 - e^{-x/\theta}$, $x > 0$, then

$$F_Y(y) = 1 - e^{-y/\theta} = W_{0,\theta}(y), \quad y > 0$$

- **"memoryless property"** the exponential: the excess of the loss over the threshold has the same distribution as the original loss r.v.

Stability of excesses of the **Pareto**

If $X \sim Pareto(\alpha, \theta)$, $F_X(x) = W_{1,\alpha,\theta}(x) = 1 - \left(\frac{x+\theta}{\theta}\right)^{-\alpha}$, $x > 0$, $\alpha, \theta > 0$, then

$$F_Y(y) = 1 - \left(\frac{y + (d+\theta)}{\theta}\right)^{-\alpha} = W_{1,\alpha,d+\theta}(y), \quad y > 0$$

- The excess over the threshold of a Pareto distribution has iftself a Pareto distribution that is the same as the original loss random variable, but with a change of scale from θ to $d + \theta$.