Chapter 4

Location-Scale-Based Parametric Distributions

William Q. Meeker and Luis A. Escobar
Iowa State University and Louisiana State University

Based on the authors' text *Statistical Methods for Reliability Data,* John Wiley & Sons Inc. 1998.

July 30, 2009
8h 3min

Chapter 4
Location-Scale-Based Parametric Distributions
Objectives

- Explain importance of parametric models in the analysis of reliability data.
- Define important functions of model parameter that are of interest in reliability studies.
- Introduce the location-scale and log-location-scale families of distributions.
- Describe the properties of the exponential distribution.
- Describe the Weibull and lognormal distributions and the related underlying location-scale distributions.

Motivation for Parametric Models

- Complements nonparametric techniques.
- Parametric models can be described concisely with just a few parameters, instead of having to report an entire curve.
- It is possible to use a parametric model to extrapolate (in time) to the lower or upper tail of a distribution.
- Parametric models provide smooth estimates of failure-time distributions.

In practice it is often useful to compare various parametric and nonparametric analyses of a data set.

Functions of the Parameters

- Cumulative distribution function (cdf) of T

$$F(t;\boldsymbol{\theta}) = \Pr(T \le t), \quad t > 0.$$

- The p quantile of T is the smallest value t_p such that

$$F(t_p;\boldsymbol{\theta}) \ge p.$$

- Hazard function of T

$$h(t;\boldsymbol{\theta}) = \frac{f(t;\boldsymbol{\theta})}{1 - F(t;\boldsymbol{\theta})}, \quad t > 0.$$

Functions of the Parameters-Continued

- The mean time to failure, MTTF, of T (also known as expectation of T)

$$\mathrm{E}(T) = \int_0^\infty t f(t;\boldsymbol{\theta})\, dt = \int_0^\infty [1 - F(t;\boldsymbol{\theta})]\, dt.$$

If $\int_0^\infty t f(t;\boldsymbol{\theta})\, dt = \infty$, we say that the mean of T does **not** exist.

- The variance (or the second central moment) of T and the standard deviation

$$\begin{aligned} \mathrm{Var}(T) &= \int_0^\infty [t - \mathrm{E}(T)]^2 f(t;\boldsymbol{\theta})\, dt \\ \mathrm{SD}(T) &= \sqrt{\mathrm{Var}(T)}. \end{aligned}$$

- Coefficient of variation γ_2

$$\gamma_2 = \frac{\mathrm{SD}(T)}{\mathrm{E}(T)}.$$

Location-Scale Distributions

Y belongs to the location-scale family of distributions if the cdf of Y can be expressed as

$$F(y;\mu,\sigma) = \Pr(Y \le y) = \Phi\left(\frac{y-\mu}{\sigma}\right), \quad -\infty < y < \infty$$

where $-\infty < \mu < \infty$ is a location parameter and $\sigma > 0$ is a scale parameter.

Φ is the cdf of Y when $\mu = 0$ and $\sigma = 1$ and Φ does not depend on any unknown parameters.

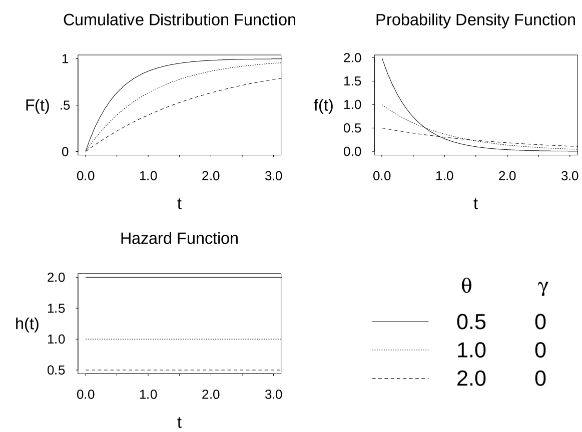
Note: The distribution of $Z = (Y - \mu)/\sigma$ does **not** depend on any unknown parameters.

Importance of Location-Scale Distributions

Importance due to:

- Most widely used statistical distributions are either members of this class or closely related to this class of distributions: exponential, normal, Weibull, lognormal, loglogistic, logistic, and extreme value distributions.
- Methods of inference, statistical theory, and computer software generated for the general family can be applied to this large, important class of models.
- Theory for location-scale distributions is relatively simple.

Examples of Exponential Distributions



Exponential Distribution

For $T \sim \text{EXP}(\theta, \gamma)$,

$$\begin{aligned} F(t;\theta,\gamma) &= 1 - \exp\left(-\frac{t-\gamma}{\theta}\right) \\ f(t;\theta,\gamma) &= \frac{1}{\theta}\exp\left(-\frac{t-\gamma}{\theta}\right) \\ h(t;\theta,\gamma) &= \frac{f(t;\theta,\gamma)}{1-F(t;\theta,\gamma)} = \frac{1}{\theta}, \quad t > \gamma, \end{aligned}$$

where $\theta > 0$ is a scale parameter and γ is both a location and a threshold parameter. When $\gamma = 0$ one gets the well-known one-parameter exponential distribution.

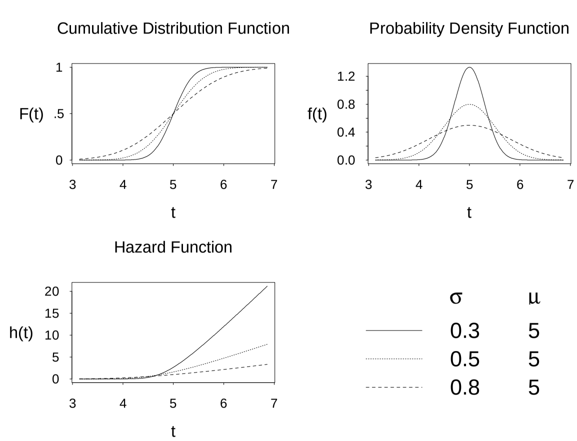
Quantiles: $t_p = \gamma - \theta \log(1-p)$.
Moments: For integer $m > 0$, $\text{E}[(T-\gamma)^m] = m!\,\theta^m$. Then

$$\text{E}(T) = \gamma + \theta, \quad \text{Var}(T) = \theta^2.$$

Motivation for the Exponential Distribution

- Simplest distribution used in the analysis of reliability data.
- Has the important characteristic that its hazard function is constant (does not depend on time t).
- Popular distribution for some kinds of electronic components (e.g., capacitors or robust, high-quality integrated circuits).
- This distribution would *not* be appropriate for a population of electronic components having failure-causing quality-defects.
- Might be useful to describe failure times for components that exhibit physical wearout only after expected technological life of the system in which the component would be installed.

Examples of Normal Distributions



Normal (Gaussian) Distribution

For $Y \sim \text{NOR}(\mu, \sigma)$

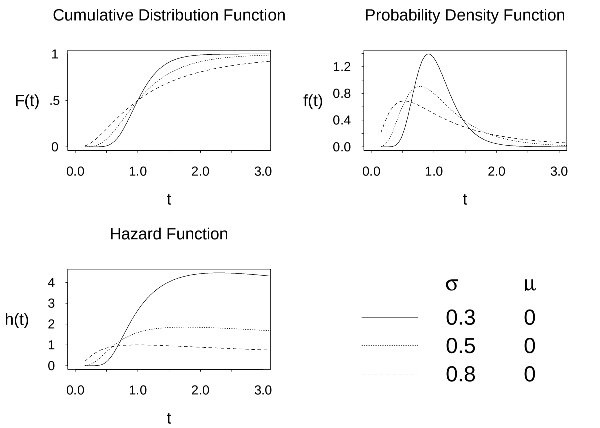
$$\begin{aligned} F(y;\mu,\sigma) &= \Phi_{\text{nor}}\left(\frac{y-\mu}{\sigma}\right) \\ f(y;\mu,\sigma) &= \frac{1}{\sigma}\phi_{\text{nor}}\left(\frac{y-\mu}{\sigma}\right), \quad -\infty < y < \infty. \end{aligned}$$

where $\phi_{\text{nor}}(z) = (1/\sqrt{2\pi})\exp(-z^2/2)$ and $\Phi_{\text{nor}}(z) = \int_{-\infty}^{z} \phi_{\text{nor}}(w)dw$ are pdf and cdf for a standardized normal ($\mu = 0, \sigma = 1$). $-\infty < \mu < \infty$ is a location parameter; $\sigma > 0$ is a scale parameter.

Quantiles: $y_p = \mu + \sigma\Phi_{\text{nor}}^{-1}(p)$ where $\Phi_{\text{nor}}^{-1}(p)$ is the p quantile for a standardized normal.
Moments: For integer $m > 0$, $E[(Y-\mu)^m] = 0$ if m is odd, and $E[(Y-\mu)^m] = (m)!\sigma^m/[2^{m/2}(m/2)!]$ if m is even. Thus

$$\text{E}(Y) = \mu, \quad \text{Var}(Y) = \sigma^2.$$

Examples of Lognormal Distributions



Lognormal Distribution

If $T \sim$ LOGNOR(μ, σ) then $\log(T) \sim$ NOR(μ, σ) with

$$F(t; \mu, \sigma) = \Phi_{\text{nor}}\left[\frac{\log(t) - \mu}{\sigma}\right]$$

$$f(t; \mu, \sigma) = \frac{1}{\sigma t}\phi_{\text{nor}}\left[\frac{\log(t) - \mu}{\sigma}\right], \quad t > 0.$$

ϕ_{nor} and Φ_{nor} are pdf and cdf for a standardized normal. $\exp(\mu)$ is a scale parameter; $\sigma > 0$ is a shape parameter.

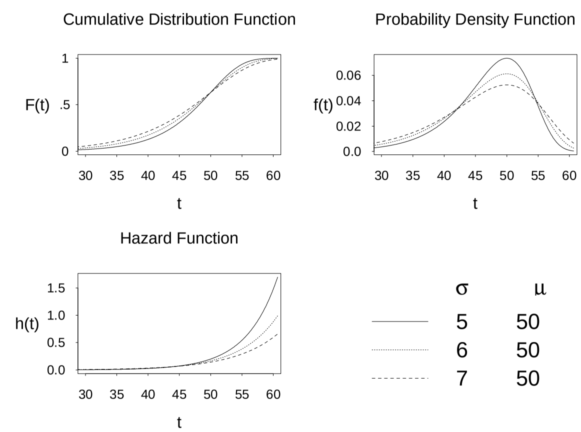
Quantiles: $t_p = \exp\left(\mu + \sigma\Phi_{\text{nor}}^{-1}(p)\right)$, where $\Phi_{\text{nor}}^{-1}(p)$ is the p quantile for a standardized normal.
Moments: For integer $m > 0$, $E(T^m) = \exp\left(m\mu + m^2\sigma^2/2\right)$.

$\text{E}(T) = \exp\left(\mu + \sigma^2/2\right)$, $\text{Var}(T) = \exp\left(2\mu + \sigma^2\right)\left[\exp(\sigma^2) - 1\right]$.

Motivation for Lognormal Distribution

- The lognormal distribution is a common model for failure times.
- It can be justified for a random variable that arises from the product of a number of identically distributed independent positive random quantities.
- It has been suggested as an appropriate model for failure time caused by a degradation process with combinations of random rates that combine multiplicatively.
- Widely used to describe time to fracture from fatigue crack growth in metals.
- Useful in modeling failure time of a population electronic components with a decreasing hazard function (due to a small proportion of defects in the population).
- Useful for describing the failure-time distribution of certain degradation processes.

Examples of Smallest Extreme Value Distributions



Smallest Extreme Value Distribution

For $Y \sim$ SEV(μ, σ),

$$F(y; \mu, \sigma) = \Phi_{\text{sev}}\left(\frac{y - \mu}{\sigma}\right)$$

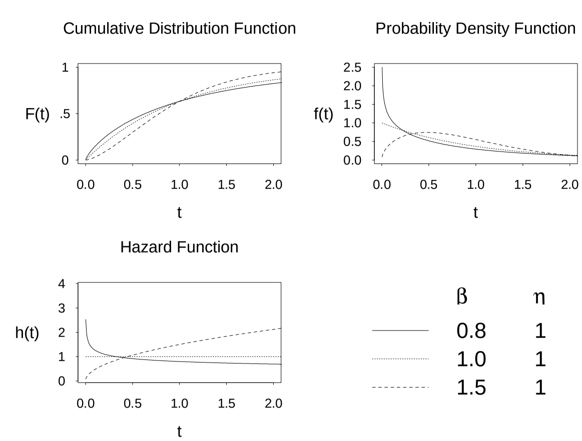
$$f(y; \mu, \sigma) = \frac{1}{\sigma}\phi_{\text{sev}}\left(\frac{y - \mu}{\sigma}\right)$$

$$h(y; \mu, \sigma) = \frac{1}{\sigma}\exp\left(\frac{y - \mu}{\sigma}\right), \quad -\infty < y < \infty.$$

$\Phi_{\text{sev}}(z) = 1 - \exp[-\exp(z)]$, $\phi_{\text{sev}}(z) = \exp[z - \exp(z)]$ are cdf and pdf for standardized SEV ($\mu = 0, \sigma = 1$). $-\infty < \mu < \infty$ is a location parameter and $\sigma > 0$ is a scale parameter.

Quantiles: $y_p = \mu + \Phi_{\text{sev}}^{-1}(p)\sigma = \mu + \log\left[-\log(1 - p)\right]\sigma$.
Mean and Variance: $\text{E}(Y) = \mu - \sigma\gamma$, $\text{Var}(Y) = \sigma^2\pi^2/6$, where $\gamma \approx .5772, \pi \approx 3.1416$.

Examples of Weibull Distributions



Weibull Distribution

Common Parameterization:

$$F(t) = \Pr(T \le t) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right]$$
$$f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}\exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right]$$
$$h(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}, \quad t > 0$$

$\beta > 0$ is shape parameter; $\eta > 0$ is scale parameter.

Quantiles: $t_p = \eta\,[-\log(1-p)]^{1/\beta}$.
Moments: For integer $m > 0$, $E(T^m) = \eta^m \Gamma(1+m/\beta)$. Then

$$E(T) = \eta\Gamma\left(1+\frac{1}{\beta}\right), \quad Var(T) = \eta^2\left[\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma^2\left(1+\frac{1}{\beta}\right)\right]$$

where $\Gamma(\kappa) = \int_0^\infty w^{\kappa-1}\exp(-w)dw$ is the gamma function.

Note: When $\beta = 1$ then $T \sim EXP(\eta)$.

Alternative Weibull Parameterization

Note: If $T \sim WEIB(\mu,\sigma)$ then $Y = \log(T) \sim SEV(\mu,\sigma)$.

For $T \sim WEIB(\mu,\sigma)$ then

$$F(t;\mu,\sigma) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] = \Phi_{sev}\left[\frac{\log(t)-\mu}{\sigma}\right]$$
$$f(t;\mu,\sigma) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}\exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] = \frac{1}{\sigma t}\phi_{sev}\left[\frac{\log(t)-\mu}{\sigma}\right]$$

where $\sigma = 1/\beta$, $\mu = \log(\eta)$, and

$$\phi_{sev}(z) = \exp[z - \exp(z)]$$
$$\Phi_{sev}(z) = 1 - \exp[-\exp(z)].$$

Quantiles:

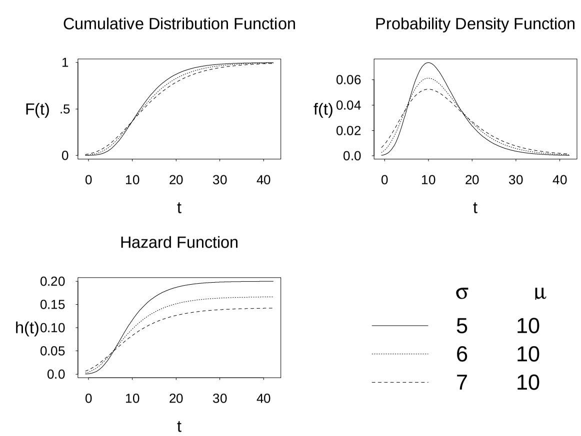
$$t_p = \eta\,[-\log(1-p)]^{1/\beta} = \exp\left[\mu + \sigma\Phi^{-1}_{sev}(p)\right]$$

where $\Phi^{-1}_{sev}(p)$ is the p quantile for a standardized SEV (i.e., $\mu = 0, \sigma = 1$).

Motivation for the Weibull Distribution

- The theory of extreme values shows that the Weibull distribution can be used to model the minimum of a large number of independent positive random variables from a certain class of distributions.
  - Failure of the weakest link in a chain with many links with failure mechanisms (e.g., creep or fatigue) in each link acting approximately independent.
  - Failure of a system with a large number of components in series and with approximately independent failure mechanisms in each component.
- The more common justification for its use is empirical: the Weibull distribution can be used to model failure-time data with a decreasing or an increasing hazard function.

Examples of Largest Extreme Value Distributions



Largest Extreme Value Distribution

When $Y \sim LEV(\mu,\sigma)$,

$$F(y;\mu,\sigma) = \Phi_{lev}\left(\frac{y-\mu}{\sigma}\right)$$
$$f(y;\mu,\sigma) = \frac{1}{\sigma}\phi_{lev}\left(\frac{y-\mu}{\sigma}\right)$$
$$h(y;\mu,\sigma) = \frac{\exp\left(-\frac{y-\mu}{\sigma}\right)}{\sigma\left\{\exp\left[\exp\left(-\frac{y-\mu}{\sigma}\right)\right]-1\right\}}, \quad -\infty < y < \infty.$$

where $\Phi_{lev}(z) = \exp[-\exp(-z)]$ and $\phi_{lev}(z) = \exp[-z-\exp(-z)]$ are the cdf and pdf for a standardized LEV ($\mu = 0, \sigma = 1$) distribution.
$-\infty < \mu < \infty$ is a location parameter and $\sigma > 0$ is a scale parameter.

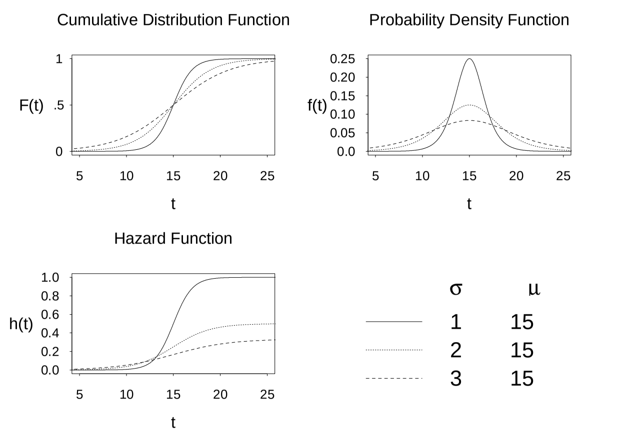
Largest Extreme Value Distribution - Continued

Quantiles: $y_p = \mu - \sigma\log[-\log(p)]$.
Mean and Variance: $E(Y) = \mu + \sigma\gamma$, $Var(Y) = \sigma^2\pi^2/6$, where $\gamma \approx .5772, \pi \approx 3.1416$.

Notes:

- The hazard is increasing but is bounded in the sense that $\lim_{y\to\infty} h(y;\mu,\sigma) = 1/\sigma$.
- If $Y \sim LEV(\mu,\sigma)$ then $-Y \sim SEV(-\mu,\sigma)$.

Examples of Logistic Distributions



Logistic Distribution

For $Y \sim \text{LOGIS}(\mu, \sigma)$,

$$\begin{aligned} F(y; \mu, \sigma) &= \Phi_{\text{logis}}\left(\frac{y-\mu}{\sigma}\right) \\ f(y; \mu, \sigma) &= \frac{1}{\sigma}\phi_{\text{logis}}\left(\frac{y-\mu}{\sigma}\right) \\ h(y; \mu, \sigma) &= \frac{1}{\sigma}\Phi_{\text{logis}}\left(\frac{y-\mu}{\sigma}\right), \quad -\infty < y < \infty. \end{aligned}$$

$-\infty < \mu < \infty$ is a location parameter; $\sigma > 0$ is a scale parameter.

ϕ_{logis} and Φ_{logis} are pdf and cdf for a standardized logistic distribution defined by

$$\begin{aligned} \phi_{\text{logis}}(z) &= \frac{\exp(z)}{[1+\exp(z)]^2} \\ \Phi_{\text{logis}}(z) &= \frac{\exp(z)}{1+\exp(z)}. \end{aligned}$$

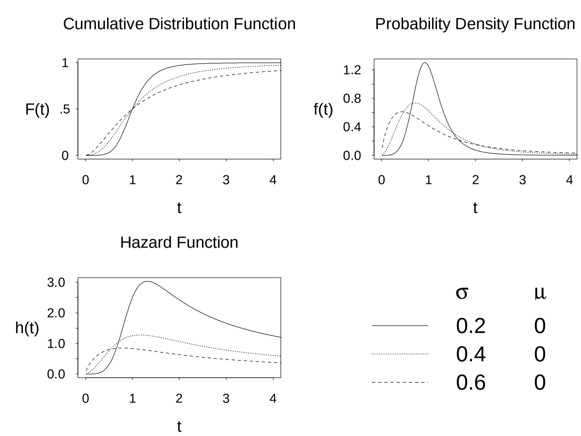
Logistic Distribution-Continued

Quantiles: $y_p = \mu + \sigma\Phi^{-1}_{\text{logis}}(p) = \mu + \sigma \log\left(\frac{p}{1-p}\right)$, where $\Phi^{-1}_{\text{logis}}(p) = \log[p/(1-p)]$ is the p quantile for a standardized logistic distribution.

Moments: For integer $m > 0$, $E[(Y-\mu)^m] = 0$ if m is odd, and $E[(Y-\mu)^m] = 2\sigma^m\,(m!)\left[1-(1/2)^{m-1}\right]\sum_{i=1}^{\infty}(1/i)^m$ if m is even. Thus

$$\text{E}(Y) = \mu, \quad \text{Var}(Y) = \frac{\sigma^2\pi^2}{3}.$$

Examples of Loglogistic Distributions



Loglogistic Distribution

If $Y \sim \text{LOGIS}(\mu, \sigma)$ then $T = \exp(Y) \sim \text{LOGLOGIS}(\mu, \sigma)$ with

$$\begin{aligned} F(t; \mu, \sigma) &= \Phi_{\text{logis}}\left[\frac{\log(t)-\mu}{\sigma}\right] \\ f(t; \mu, \sigma) &= \frac{1}{\sigma t}\phi_{\text{logis}}\left[\frac{\log(t)-\mu}{\sigma}\right] \\ h(t; \mu, \sigma) &= \frac{1}{\sigma t}\Phi_{\text{logis}}\left[\frac{\log(t)-\mu}{\sigma}\right], \quad t > 0. \end{aligned}$$

$\exp(\mu)$ is a scale parameter; $\sigma > 0$ is a shape parameter. Φ_{logis} and ϕ_{logis} are cdf and pdf for a $\text{LOGIS}(0, 1)$.

Loglogistic Distribution-Continued

Quantiles: $t_p = \exp\left[\mu + \sigma\Phi^{-1}_{\text{logis}}(p)\right] = \exp(\mu)\,[p/(1-p)]^{\sigma}$.
Moments: For integer $m > 0$,

$$\text{E}(T^m) = \exp(m\mu)\,\Gamma(1+m\sigma)\,\Gamma(1-m\sigma).$$

The m moment is not finite when $m\sigma \geq 1$.

For $\sigma < 1$,

$$\text{E}(T) = \exp(\mu)\,\Gamma(1+\sigma)\,\Gamma(1-\sigma),$$

and for $\sigma < 1/2$,

$$\text{Var}(T) = \exp(2\mu)\,\left[\Gamma(1+2\sigma)\,\Gamma(1-2\sigma) - \Gamma^2(1+\sigma)\,\Gamma^2(1-\sigma)\right].$$

Other Topics in Chapter 4

Pseudorandom number generation.

Topics in Chapter 5

- Parametric models with threshold parameters.
- Important distributions used in reliability that can not be translated into location-scale distributions: gamma, generalized gamma, etc.
- Finite (discrete) mixture distributions

$$F(t;\boldsymbol{\theta}) = \xi_1 F_1(t;\boldsymbol{\theta}_1) + \cdots + \xi_k F_k(t;\boldsymbol{\theta}_k)$$

  where $\xi_i \geq 0$, and $\sum_i \xi_i = 1$
- Compound (continuous) mixture distributions.

  If failure-times of units in a population are EXP(η) with $1/\eta \sim$ GAM(θ, κ), then the unconditional failure time, T, of a unit selected at random from the population has a Pareto distribution of the form

$$F(t;\theta,\kappa) = 1 - \frac{1}{(1+\theta t)^{\kappa}}, \quad t > 0.$$