

Unit 4: Location-Scale-Based Parametric Distributions

Ramón V. León

Notes largely based on
“Statistical Methods for Reliability Data”
by W.Q. Meeker and L. A. Escobar,
Wiley, 1998 and on their class notes.

Unit 4 Objectives

- Explain the importance of parametric models in the analysis of reliability data
- Define important functions of model parameters that are of interest in reliability studies
- Introduce the location-scale family 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

- Complement 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 analysis of a data set.

Function of the Parameters

- Cumulative distribution function (cdf) of T

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

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

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

- 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)

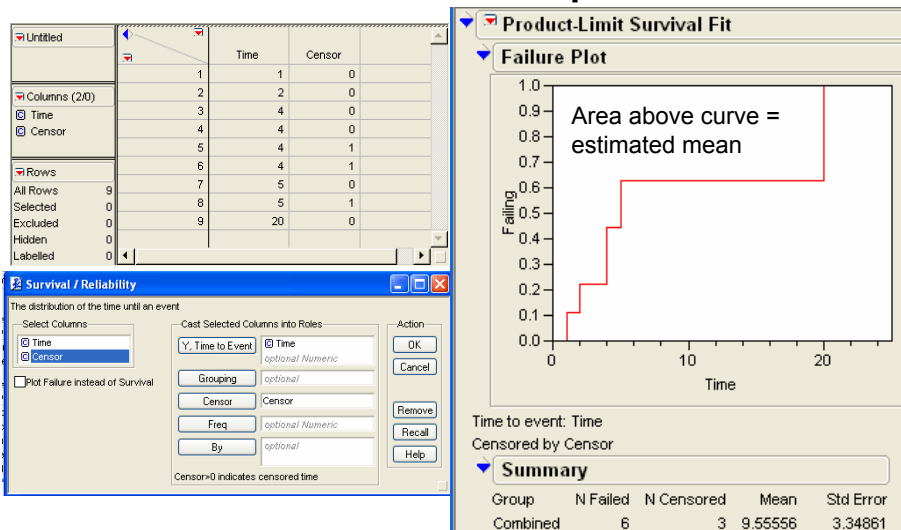
$$E(T) = \int_0^{\infty} t f(t; \theta) dt = \int_0^{\infty} [1 - F(t; \theta)] dt.$$

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

Remark:
$$\hat{\mu} = \int_0^{\infty} [1 - \hat{F}(t)] dt = \int_0^{\infty} [\hat{S}(t)] dt$$

if the last time is a failure time so that $\hat{S}(t)$ reaches 0 at that time.

JMP Example



Functions of the Parameters-Continued

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

$$\text{Var}(T) = \int_0^{\infty} [t - E(T)]^2 f(t; \theta) dt$$

$$\text{SD}(T) = \sqrt{\text{Var}(T)}.$$

- Coefficient of variation γ_2

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

Location-Scale Distributions

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

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

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

- 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 relative simple

One Parameter Exponential Distribution Parametrized by the Hazard Rate

$$f(t) = \lambda e^{-\lambda t}, \quad F(t) = \int_0^t \lambda e^{-\lambda x} dx = 1 - e^{-\lambda t},$$

$$S(t) = e^{-\lambda t}, \quad h(t) = \lambda \text{ for } t \geq 0$$

$$E(T) = \frac{1}{\lambda} \text{ and } Var(T) = \frac{1}{\lambda^2}$$

Two-Parameter Exponential Distribution

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

$$F(t; \theta, \gamma) = 1 - \exp\left(-\frac{t-\gamma}{\theta}\right)$$

$$f(t; \theta, \gamma) = \frac{1}{\theta} \exp\left(-\frac{t-\gamma}{\theta}\right)$$

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

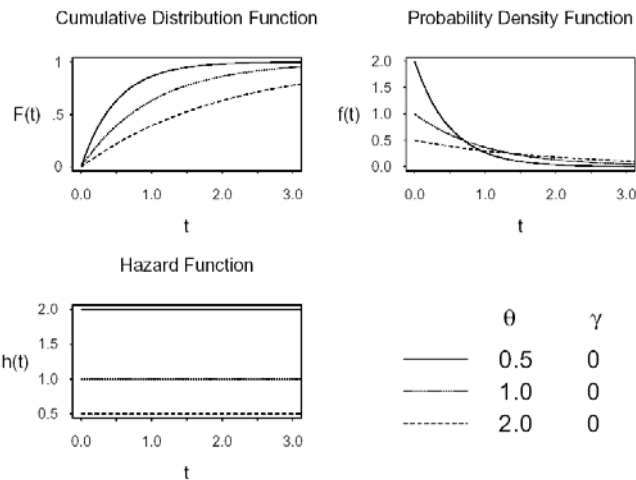
where $\theta > 0$ is a scale parameter and γ is both a location and a threshold parameter. When $\gamma = 0$ one gets the well-known one-parameter exponential distribution.

Quantiles: $t_p = \gamma - \theta \log(1 - p)$.

Moments: For integer $m > 0$, $E[(T - \gamma)^m] = m! \theta^m$. Then

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

Examples of Exponential Distributions



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

Normal (Gaussian) Distribution

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

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

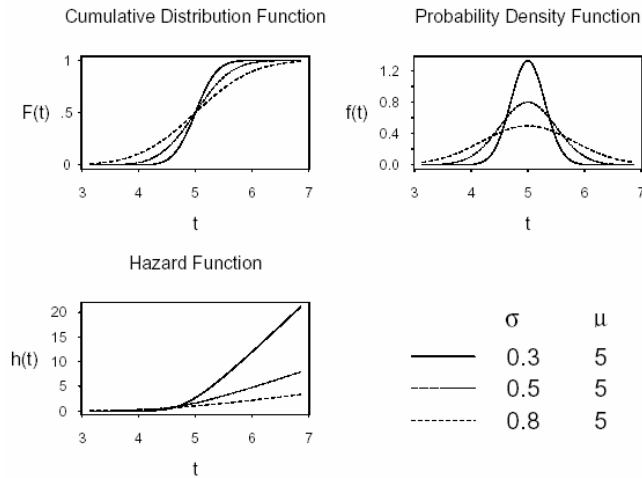
where $\phi_{\text{nor}}(z) = (1/\sqrt{2\pi}) \exp(-z^2/2)$ and $\Phi_{\text{nor}}(z) = \int_{-\infty}^z \phi_{\text{nor}}(w) dw$ are pdf and cdf for a standardized normal ($\mu = 0, \sigma = 1$).
 $-\infty < \mu < \infty$ is a location parameter; $\sigma > 0$ is a scale parameter.

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

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

Examples of Normal Distributions



8/31/2004

Stat 567: Unit 4 - Ramón V. León

15

Lognormal Distribution

If $T \sim \text{LOGNOR}(\mu, \sigma)$ then $\log(T) \sim \text{NOR}(\mu, \sigma)$ 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: $t_p = \exp(\mu + \sigma \Phi_{\text{nor}}^{-1}(p))$, where $\Phi_{\text{nor}}^{-1}(p)$ is the p quantile for a standardized normal.

Moments: For integer $m > 0$, $E(T^m) = \exp(m\mu + m^2\sigma^2/2)$.

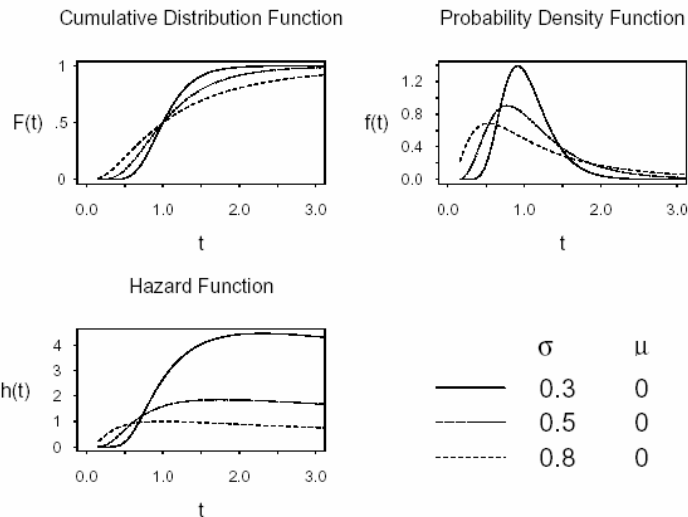
$$E(T) = \exp(\mu + \sigma^2/2), \quad \text{Var}(T) = \exp(2\mu + \sigma^2) [\exp(\sigma^2) - 1].$$

8/31/2004

Stat 567: Unit 4 - Ramón V. León

16

Examples of Lognormal Distributions



8/31/2004

Stat 567: Unit 4 - Ramón V. León

17

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 a product of a number of identically distributed independent positive random quantities
- It has been suggested as an appropriate model for failure times 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 of 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

8/31/2004

Stat 567: Unit 4 - Ramón V. León

18

Smallest Extreme Value (Gumbel) Distribution

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

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

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

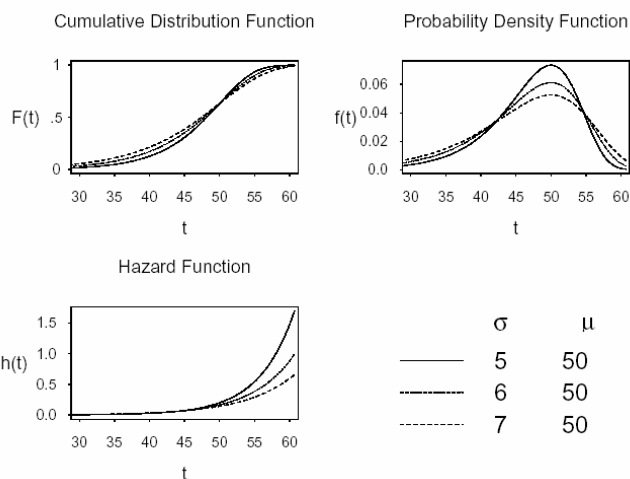
$$h(y; \mu, \sigma) = \frac{1}{\sigma} \exp\left(\frac{y - \mu}{\sigma}\right), \quad -\infty < y < \infty.$$

$\Phi_{\text{sev}}(z) = 1 - \exp[-\exp(z)]$, $\phi_{\text{sev}}(z) = \exp[z - \exp(z)]$ 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: $y_p = \mu + \Phi_{\text{sev}}^{-1}(p)\sigma = \mu + \log[-\log(1 - p)]\sigma$.

Mean and Variance: $E(Y) = \mu - \sigma\gamma$, $\text{Var}(Y) = \sigma^2\pi^2/6$, where $\gamma \approx .5772$, $\pi \approx 3.1416$.

Examples of Smallest Extreme Value Distributions



Weibull Distribution

Common Parameterization:

$$F(t) = \Pr(T \leq t) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^\beta\right]$$

$$f(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\eta}\right)^\beta\right]$$

$$h(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}, \quad t > 0$$

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

Quantiles: $t_p = \eta[-\log(1-p)]^{1/\beta}$.

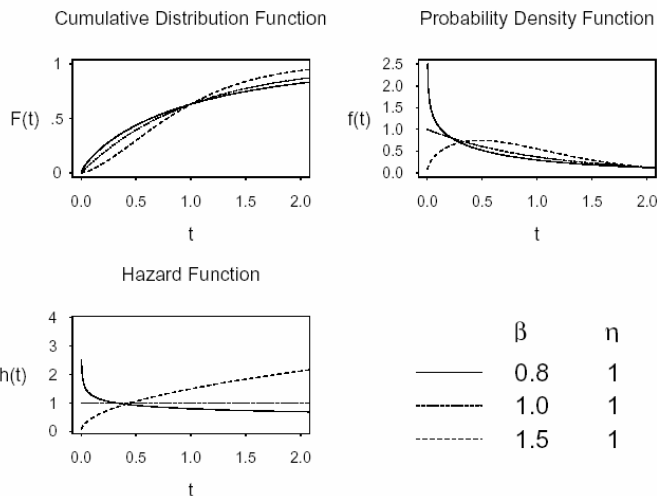
Moments: For integer $m > 0$, $E(T^m) = \eta^m \Gamma(1+m/\beta)$. Then

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

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

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

Examples of Weibull Distributions



Alternative Weibull Parametrization

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

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

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

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

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

Quantiles:

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

where $\Phi_{\text{sev}}^{-1}(p)$ 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 rate

Largest Extreme Value Distribution

When $Y \sim \text{LEV}(\mu, \sigma)$,

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

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

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

where $\Phi_{\text{lev}}(z) = \exp[-\exp(-z)]$ and $\phi_{\text{lev}}(z) = \exp[-z - \exp(-z)]$ 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

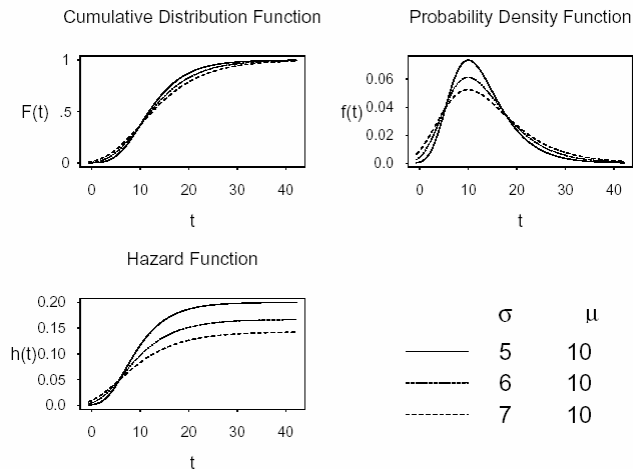
Quantiles: $y_p = \mu - \sigma \log[-\log(p)]$.

Mean and Variance: $E(Y) = \mu + \sigma\gamma$, $\text{Var}(Y) = \sigma^2\pi^2/6$,
where $\gamma \approx .5772$, $\pi \approx 3.1416$.

Notes:

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

Examples of the Largest Extreme Value Distribution



8/31/2004

Stat 567: Unit 4 - Ramón V. León

27

Logistic Distribution

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

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

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

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

$-\infty < \mu < \infty$ 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

$$\phi_{\text{logis}}(z) = \frac{\exp(z)}{[1 + \exp(z)]^2}$$

$$\Phi_{\text{logis}}(z) = \frac{\exp(z)}{1 + \exp(z)}.$$

8/31/2004

Stat 567: Unit 4 - Ramón V. León

28

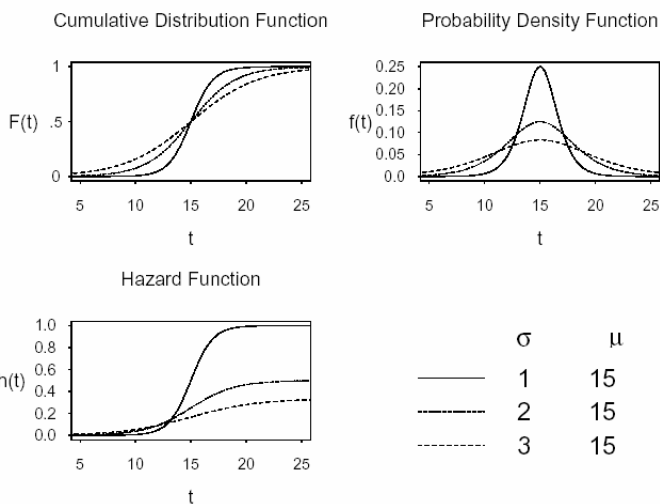
Logistic Distribution - Continued

Quantiles: $y_p = \mu + \sigma \Phi_{\text{logis}}^{-1}(p) = \mu + \sigma \log\left(\frac{p}{1-p}\right)$, where $\Phi_{\text{logis}}^{-1}(p) = \log[p/(1-p)]$ 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

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

Examples of Logistic 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: $t_p = \exp \left[\mu + \sigma \Phi_{\text{logis}}^{-1}(p) \right] = \exp(\mu) [p/(1-p)]^\sigma$.

Moments: For integer $m > 0$,

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

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

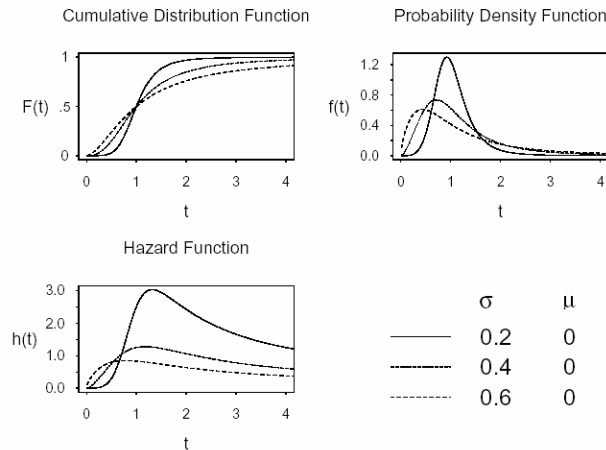
For $\sigma < 1$,

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

and for $\sigma < 1/2$,

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

Examples of Loglogistic Distributions



8/31/2004

Stat 567: Unit 4 - Ramón V. León

33

Other Topics in Chapter 4

- Pseudorandom number generation
 - Efficient method for dealing with random samples involving censoring.

8/31/2004

Stat 567: Unit 4 - Ramón V. León

34