TMA4275 LIFETIME ANALYSIS

Slides 4: Gumbel distribution. Log-location-scale families

Bo Lindqvist
Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim

http://www.math.ntnu.no/∼bo/
bo@math.ntnu.no

NTNU, Spring 2015

Let $T_1, T_2, \cdots, T_n$ be lifetimes of n components, with ordered values denoted by $T_{(1)} < T_{(2)} < \cdots < T_{(n)}$. Thus $T_{(1)}$ is the minimum and corresponds to the lifetime of a series system.

For large n, $T_{(1)}$ is approximately Weibull-distributed. *This motivates the widespread use of the Weibull-distribution!*

If the T_i are no longer lifetimes, but have support in $(-\infty, \infty)$, then the limiting distribution of a properly normalized version of $T_{(1)}$ equals the distribution of a random variable Y with cdf

$$F_Y(y) = 1 - e^{-e^{\frac{y-\mu}{\sigma}}}, \qquad -\infty < y < \infty$$

This is the so called "Distribution of smallest extreme", or "Extreme value distribution of type I", or (which we will call it) the **Gumbel-distribution**; $Y \sim \text{Gumbel}(\mu, \sigma)$.

We write $Y \sim \text{Gumbel}(\mu, \sigma)$

EXERCISE

Show that

$$F_Y(y) = 1 - e^{-e^{\frac{y-\mu}{\sigma}}}, \qquad -\infty < y < \infty$$

satisfies the requirements for a cdf, i.e.

- Increasing in y
- $\lim_{y \to -\infty} F_Y(y) = 0$
- $\lim_{y \to \infty} F_Y(y) = 1$

WHY ARE WE INTERESTED IN THE GUMBEL DISTRIBUTION?

If T is Weibull-distributed, $T \sim Weib(\alpha, \theta)$, then $Y = \ln T$ is Gumbel-distributed, $Y \sim Gumbel(\mu, \sigma)$, with $\mu = \ln\theta$, $\sigma = 1/\alpha$.

Proof: Note first that $T = e^Y$ and $R(t) = P(T > t) = e^{-\left(\frac{t}{\theta}\right)^\alpha}$. Then:

$$
\begin{aligned}
P(Y > y) &= P(e^Y > e^y) = P(T > e^y) = R(e^y) \\
&= e^{-\left(\frac{e^y}{\theta}\right)^\alpha} = e^{-\left(\frac{e^y}{e^{\ln\theta}}\right)^\alpha} \\
&= e^{-(e^{y-\ln\theta})^\alpha} = e^{-e^{\left(\frac{y-\ln\theta}{1/\alpha}\right)}}
\end{aligned}
$$

Thus, $F_Y(y) = 1 - P(Y > y) = 1 - e^{-e^{\left(\frac{y-\ln\theta}{1/\alpha}\right)}}$, which shows that $Y \sim \text{Gumbel}(\ln\theta, 1/\alpha)$.

We shall see later why this is a useful and interesting result (and not just a curiosity...)

Let $Y \sim \text{Gumbel}(\mu, \sigma)$ and recall the cdf

$$F_Y(y) = P(Y \leq y) = 1 - e^{-e^{\frac{y-\mu}{\sigma}}} \quad \text{for} \quad -\infty < y < \infty$$

The cdf of Gumbel(0,1), called *the standard Gumbel distribution*, is

$$G(w) = 1 - e^{-e^{w}} \quad \text{for} \quad -\infty < w < \infty$$

Suppose $W \sim \text{Gumbel}(0, 1)$ and let

$$Y = \mu + \sigma W$$

Then

$$F_Y(y) = P(Y \leq y) = P(\mu + \sigma W \leq y) = P(W \leq \frac{y-\mu}{\sigma}) = G\left(\frac{y-\mu}{\sigma}\right)$$

so $Y \sim \text{Gumbel}(\mu, \sigma)$.

THE GUMBEL DISTRIBUTION (CONT.)

Thus we can represent $Y \sim \text{Gumbel}(\mu, \sigma)$ as

$$Y = \mu + \sigma W$$

where $W \sim \text{Gumbel}(0, 1)$. Further,

$$F_Y(y) = P(Y \leq y) = G\left(\frac{y - \mu}{\sigma}\right)$$

where $G(\cdot)$ is the cdf of Gumbel(0,1).

This defines the cdf of the Gumbel(μ, σ) in terms of the cdf of the standard Gumbel, in the same way as the cdf of $Y \sim N(\mu, \sigma)$ can be expressed by the cdf of the standard normal.

Recall once more that if $W \sim \text{Gumbel}(0,1)$, then W has the cdf

$$G(w) = 1 - e^{-e^w}$$

The pdf of W is hence

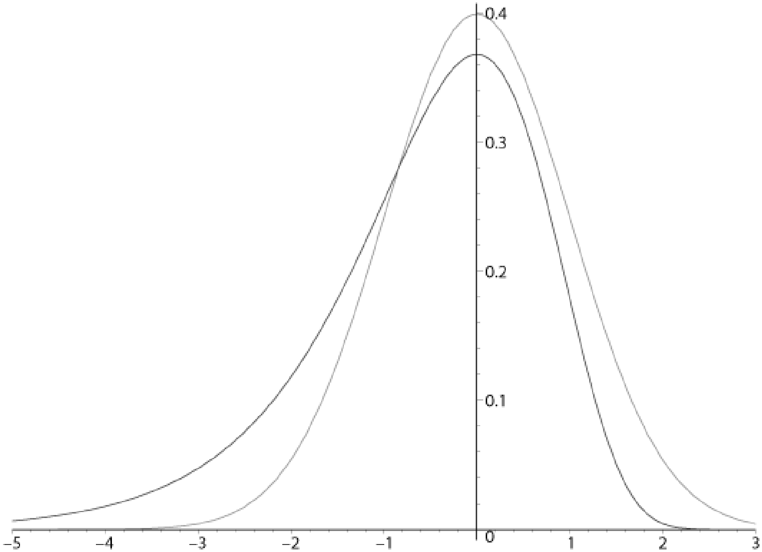
$$g(w) = G'(w) = -e^{-e^w}(-e^w) = e^w e^{-e^w}$$

We also have

$$E(W) = \int_{-\infty}^{\infty} w e^w e^{-e^w} dw = -\gamma,$$

where $\gamma = -0.5772$ is *Euler's constant*.

STANDARD GUMBEL AND NORMAL DISTRIBUTIONS



GENERAL LOG-LOCATION-SCALE FAMILIES

We have seen:

$T \sim \text{lognorm}(\mu, \sigma) \iff Y = \ln T \sim N(\mu, \sigma)$

$T \sim \text{Weib}(\alpha, \theta) \iff Y = \ln T \sim \text{Gumbel}(\mu, \sigma)$, with $\mu = \ln \theta$, $\sigma = 1/\alpha$.

- Both distributions thus define **log-location-scale families,** which are characterized by the fact that $Y = \ln T$ has a cdf which can be expressed as
$$F_Y(y) = P(Y \leq y) = \Psi\left(\frac{y-\mu}{\sigma}\right)$$
where $\Psi(\cdot)$ is the cdf of some "standardized distribution" on $(-\infty, \infty)$.
- Equivalently, log-location-scale families are characterized by representations
$$\ln T = \mu + \sigma U$$
where U has cdf $\Psi(\cdot)$ as described above.

In the representation

$$\ln T = \mu + \sigma U,$$

- U has a "standard" distribution with support $(-\infty, +\infty)$, (e.g. $N(0,1)$, Gumbel$(0,1)$)
- $\mu \in (-\infty, +\infty)$ is called the *location parameter*
- $\sigma > 0$ is called the *scale parameter*

A random variable Y has the **logistic distribution** *with location parameter μ and scale parameter σ, written $Y \sim$ logistic(μ, σ), if*

$$F_Y(y) = P(Y \le y) = H\left(\frac{y-\mu}{\sigma}\right) \quad \text{for } -\infty < y < \infty$$

where

$$H(v) = P(V \le v) = \frac{e^v}{1+e^v} \text{ for } -\infty < v < \infty$$

is the cdf of the standard logistic distribution, logistic(0,1).

A lifetime T has the **log-logistic** *distribution with location parameter μ and scale parameter σ if $Y = \ln T \sim$ logistic(μ, σ). In this case we have the representation*

$$\ln T = \mu + \sigma V$$

where $V \sim$ logistic$(0, 1)$.

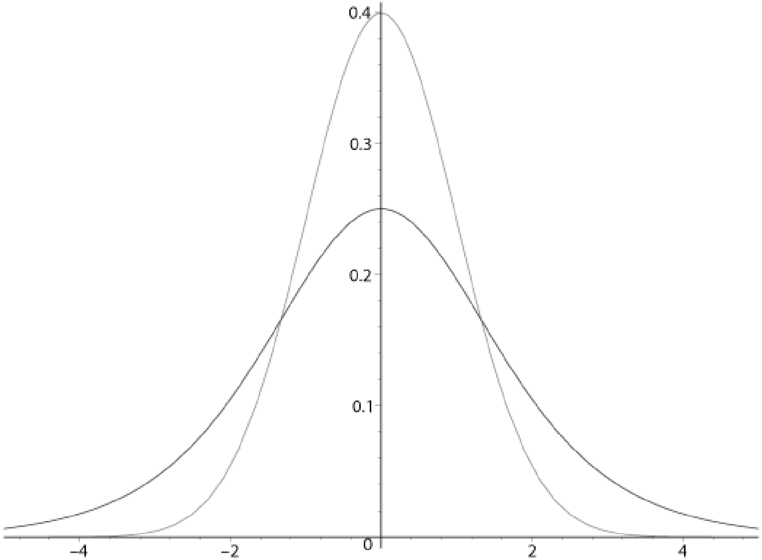
Recall that if $V \sim \text{logistic}(0, 1)$, then the cdf of V is
$H(v) = P(V \leq v) = \frac{e^v}{1+e^v}$ for $-\infty < v < \infty$.

Hence the pdf of V is

$$h(v) = H'(v) = \frac{e^v}{(1+e^v)^2} \quad \textit{(do the differentiation!)}$$

Like the standard normal, this density is symmetric around the y-axis (which is not the case for the standard Gumbel).
Check this by showing that $h(-v) = h(v)$ *for all* v.

STANDARD LOGISTIC AND STANDARD NORMAL



By assumption, $Y = \ln T$ has a cdf which can be expressed as

$$F_Y(y) = P(Y \leq y) = \Psi\left(\frac{y-\mu}{\sigma}\right) \quad \text{for } -\infty < y < \infty$$

where $\Psi(\cdot)$ is the cdf of a standard distribution. Let further $\psi(u) = \Psi'(u)$.

Then

$$\begin{aligned}
R(t) &= P(T > t) = P(\ln T > \ln t) = 1 - \Psi\left(\frac{\ln t - \mu}{\sigma}\right) \\
f(t) &= -R'(t) = \psi\left(\frac{\ln t - \mu}{\sigma}\right) \cdot \frac{1}{t\sigma} \\
z(t) &= \frac{f(t)}{R(t)} = \frac{\psi\left(\frac{\ln t - \mu}{\sigma}\right)/(t\sigma)}{1 - \Psi\left(\frac{\ln t - \mu}{\sigma}\right)}
\end{aligned}$$

(as already obtained for the lognormal distribution).

- *Extreme value distributions*
  - Weibull-distribution
  - Gumbel distribtuion, Gumbel(μ, σ)
- $T \sim$ Weibull$(\alpha, \theta) \Rightarrow \ln T \sim$ Gumbel$(\ln \theta, 1/\alpha)$
- Thus:
  - Gumbel(μ, σ) is a location-scale family
  - Weibull(α, θ) is a log-location-scale family based on Gumbel-distribution
- General definition and properties of log-location-scale families
- Another example: The *logistic* and *log-logistic* distributions