Generalized Exponential Distribution

Debasis Kundu
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur
India

Joint Work: R.D. Gupta, Univ. of New Brunswick & A. Manglick, Univ. of Umea.

Outline of the Talk

- Genesis of the Model
- Model Description
- Common Properties with Weibull and Gamma Distribution
- Moments of the GED
- Inference
- Closeness with Other Distributions
- Generation of Gamma and Normal Random Variables Using GED
- References

Genesis of the Model

Gompertz (1825, Phil. Trans. Roy. Soc. Lon.) used the following distribution function to represent mortality growth

$$G(t) = \left(1 - \rho e^{-\lambda t}\right)^{\alpha} \qquad t > \frac{1}{\lambda} \ln \rho.$$

Ahuja and Nash (1967, Sankhya A) also used this model and some related model for growth curve mortality

Gupta and Kundu (1999, ANZJS) consider a special case of this model

$$F(x; \alpha, \lambda) = \left(1 - e^{-\lambda x}\right)^{\alpha}; \qquad x > 0.$$

This is a special case of the exponentiated Weibull model proposed by Mudholkar and his co-workers (1995, Technometrics)

Model Description

Here α is the shape and λ is the scale parameter. It has different shapes of the density function

$$f(x;\alpha,\lambda) = \alpha\lambda\left(1-e^{-x\lambda}\right)^{\alpha-1} e^{-x\lambda}; \quad x > 0.$$

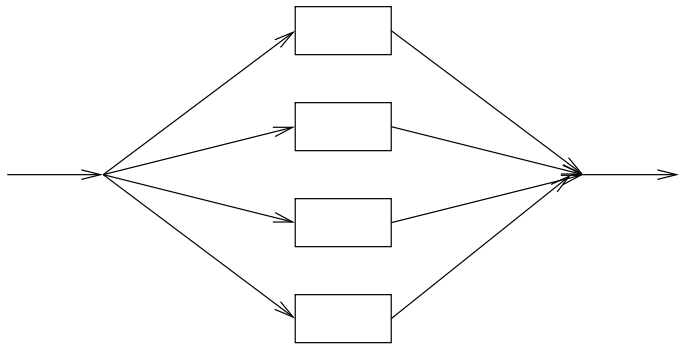
It has different shapes of the hazard functions

$$h(x;\alpha,\lambda) = \frac{f(x;\alpha,\lambda)}{1-F(x;\alpha,\lambda)} = \frac{\alpha\lambda\left(1-e^{-\lambda x}\right)^{\alpha-1} e^{-\lambda x}}{1-\left(1-e^{-\lambda x}\right)^{\alpha}}; \quad x > 0.$$

Physical Interpretations

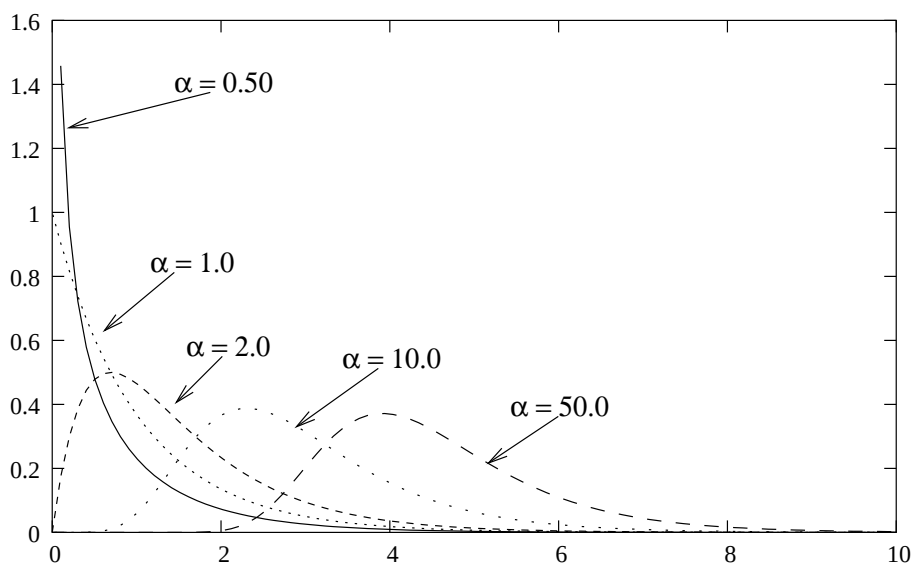
A parallel system is a system where the system works if at least one of the components works

If the shape parameter α is an integer it represents the life time of a parallel system when each component follows exponential distribution



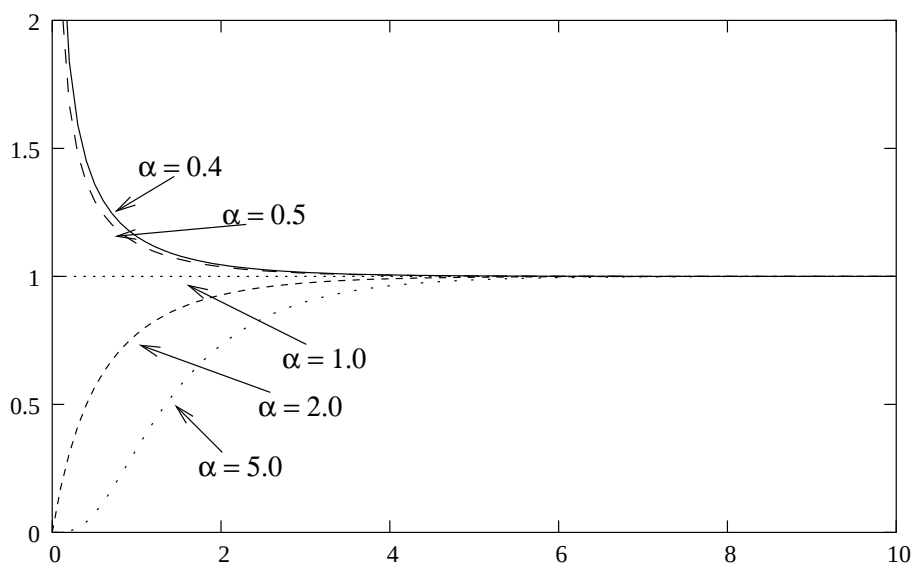
Shapes of the Different Density Functions

Since λ is the scale parameter we take $\lambda = 1$



Shapes of the Different Hazard Functions

Since λ is the scale parameter we take $\lambda = 1$



Comparisons with Gamma and Weibull Distributions

Density Functions

- Shape parameter $= 1$ all are equal to exponential distribution
- Shape parameter > 1, all are unimodal
- Shape parameter < 1, all are decreasing functions

Hazard Functions

- Shape parameter $= 1$ all have constant hazard functions
- Shape parameter > 1, all are increasing functions
- Shape parameter < 1, all are decreasing functions

Closer look at the hazard functions

Shape-Parameter	Gamma	Weibull	GE
1	1	1	1
> 1	$0 \uparrow 1$	$0 \uparrow \infty$	$0 \uparrow 1$
< 1	$\infty \downarrow 1$	$\infty \downarrow 0$	$\infty \downarrow 1$

GE is more closer to the gamma distribution rather than the Weibull distribution in this respect

Ordering Relations in Terms of the Shape Parameter

	LR	HAZ	ST
Gamma	Y	Y	Y
Weibull	N	N	N
GE	Y	Y	Y

Moments

Moment generating Function ($\lambda = 1$).

$$M(t) = \frac{\Gamma(\alpha+1)\Gamma(1-t)}{\Gamma(\alpha-t+1)} \qquad (t < 1).$$

$$E(X) = \psi(\alpha+1) - \psi(1), \qquad V(X) = \psi'(1) - \psi'(\alpha+1).$$

Stochastic representation

$$X \stackrel{d}{=} \sum_{j=1}^{[\alpha]} \frac{Y_j}{j + <\alpha>} + Z$$

Here $[\alpha]$ is the integer part of α and $<\alpha>$ is the fractional part of α. Y_j's are $i.i.d.$ exponential with mean 1 and Z follows GE with shape parameter $<\alpha>$.

$$E(X) = \sum_{i=1}^{n} \frac{1}{i}, \qquad V(X) = \sum_{i=1}^{n} \frac{1}{i^2}.$$

Sum of n $i.i.d$ GE Distribution

If $X_1, \ldots, X_n$ are n $i.i.d.$ $GE(\alpha)$, then the density function of $X = \sum_{i=1}^{n} X_i$ is

$$f_X(x) = \sum_{j=0}^{\infty} c_j f_{GE}(x; n\alpha + j)$$

Here

$$c_j > 0 \qquad \sum_{i=1}^{n} c_j = 1.$$

If we approximate it by M terms, $i.e.$

$$f_X(x) \approx \sum_{j=0}^{M-1} c_j f_{GE}(x; n\alpha + j),$$

then the error due to approximations is bounded by

$$\left(1 - \sum_{j=0}^{M-1} c_j\right) g(x),$$

Explicit expression of g is available.

Estimation

The problem is to estimate the unknown parameters from a random sample of size n.

- The family of the GED satisfy all the regularity conditions
- The MLEs work quite well if α is not very close to 0
- Fixed point type iterative process can be used to solve the non-linear equation
- If α is close to 0, the iterative process takes longer time to converge
- Other estimators like Moment Estimators, L-Moment estimators, Percentile Estimators, Least Squares Estimators, BLUE have been tried

If one of the parameters is known

- If the scale parameter is known then the MLE of the shape parameter can be obtained in explicit form
- If the shape parameter is known then the MLEs of the scale parameter can be obtained by solving a non-linear equation.
- Other estimators also can be obtained accordingly

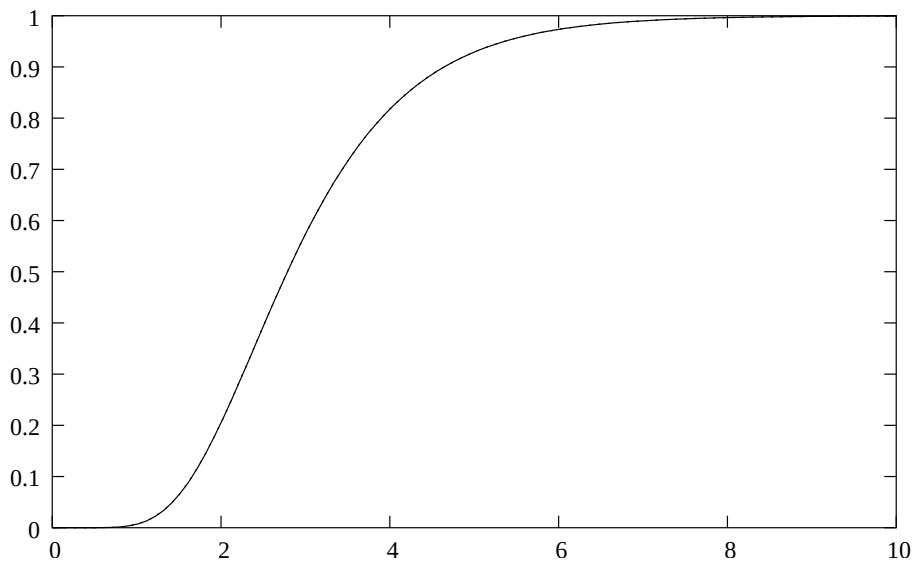
Testing and Confidence Intervals

- LRT can be used for testing purposes if both are unknown
- If λ is known it is an exponential family, then the UMP or UMPU test exists for testing the shape parameter
- If α is known then testing the scale parameter LRT can be used
- If both the parameters are unknown then asymptotic confidence intervals can be used for constructing confidence intervals
- If λ is known then exact confidence intervals based on χ^2 distribution is available

Closeness with Other Distributions

For certain ranges of the shape and scale parameters the distribution function of the GE distribution can be very close to the corresponding distribution functions of Weibull, gamma and log-normal distributions.

The distribution function of $GE(12.9)$ and LN(0.3807482, 2.9508672)



Disadvantage: It is very difficult to distinguish between the two models. Selecting the correct model for small sample sizes becomes almost impossible.

Advantage: Log-normal distribution function or gamma distribution function can be approximated very well by GE

We can generate approximate random samples of log-normal (implies normal also) and gamma distributions using GE

A very convenient approximation of the standard normal distribution function can be used as

$$\Phi(z) \approx \left(1 - e^{-e^{z\sigma+\mu}}\right)^{12.9}$$

where

$$\sigma = 0.3807482 \qquad \mu = 1.0820991$$

In this case the error of approximation is less than 0.0003.

Generation of $N(0,1)$

Using the approximation, approximate $N(0,1)$ can be easily generated

Generation of Gamma(α)

Consider the Gamma distribution with the following density function

$$f_{GA}(x;\alpha) = \frac{1}{\Gamma\alpha} x^{\alpha-1} e^{-x}; \quad x > 0$$

It can be shown that for $0 < \alpha < 1$

$$f_{GA}(x;\alpha) \leq \frac{2^{\alpha}}{\Gamma(\alpha+1)} f_{GE}(x;\alpha,\frac{1}{2})$$

Using the inequality and by Acceptance-Rejection method gamma random numbers can be generated

References

- Ahuja, J. C. and Nash, S. W. (1967), "The generalized Gompertz-Verhulst family of distributions", *Sankhya*, Ser. A., vol. 29, 141 - 156.
- Gompertz, B. (1825), "On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies", *Philosophical Transactions of the Royal Society London*, vol. 115, 513 - 585.
- Gupta, R. D. and Kundu, D. (1999). "Generalized exponential distributions", *Australian and New Zealand Journal of Statistics*, vol. 41, 173 - 188.
- Gupta, R. D. and Kundu, D. (2003), "Discriminating between the Weibull and the GE distributions", *Computational Statistics and Data Analysis*, vol. 43, 179 - 196.

References (cont.)

- Gupta, R. D. and Kundu, D. (2005), "Comparison of the Fisher information between the Weibull and generalized exponential distribution", (to appear in the *Journal of Statistical Planning and Inference*)
- Kundu, D., Gupta, R D. and Manglick, A. (2005),"Discriminating between the lognormal and generalized exponential distribution", *Journal of the Statistical Planning and Inference*, vol. 127, 213 - 227.
- Raqab, M. Z. (2002), "Inferences for generalized exponential distribution based on record statistics", *Journal of Statistical Planning and Inference*, vol. 104, 339 - 350.
- Raqab, M. Z. and Ahsanullah, M. (2001), "Estimation of the location and scale parameters of generalized exponential distribution based on order statistics", *Journal of Statistical Computation and Simulation*, vol. 69, 109 - 124.

References (cont.)

- Zheng, G. (2002), "Fisher information matrix in type -II censored data from exponentiated exponential family", *Biometrical Journal*, vol. 44, 353 - 357.