

## Estimation of Parameters of Mixed Generalized Exponentially Distributions from Censored Type I Samples

Elsherpieny. E.A

Institute of Statistical Studies & Research. Cairo University.

---

**Abstract:** In this paper a failure model which can be divided into two subpopulations, each representing generalized exponential distribution (GE) is considered. Maximum likelihood estimators (MLE) for the five unknown parameters are obtained based on type-I censored samples, as well as the elements of the information matrix, Mendenhall & Hader (1958) results may be considered as a special case from the present results. A numerical illustration is given.

---

### INTRODUCTION

Mixed failure populations are encountered in many fields of applied science. For some life testing model, it has frequently been observed that the failure rate is initially relatively high, and then actually decreases with increasing age. As the item becomes still older the failure rate either becomes constant or again increases with age depending on the basic failure mechanism involved. This behaviour suggests strongly that the population is not homogeneous but rather is made up of several subpopulations mixed in unknown proportions. For practical purpose, Mendenhall & Hader<sup>[10]</sup> mentioned that engineer may divide the failures of a system, or a device, into two or more different types of causes. An example is presented by Acheson & McElwee<sup>[11]</sup>. Other references to mixed failure populations are made in papers by Davies<sup>[3]</sup>, Epstein<sup>[4]</sup>, Steen<sup>[13]</sup>, and Wilde<sup>[14]</sup>.

Suppose we have two subpopulation represents failure types, mixed in proportion  $p_1$ ,  $(1 - p_1)$  and  $p_1 + p_2 = 1$  then the distribution function of the random variable  $X$  will be

$$F(x) = p_1 F_1(x) + p_2 F_2(x)$$

and the corresponding density function is given by

$$f(x) = p_1 f_1(x) + p_2 f_2(x)$$

Raqab & Madi<sup>[12]</sup> concluded that The two-parameter GE distribution has been introduced and studied quite extensively by Gupta and Kundu<sup>[5]</sup>. Suppose that  $n$  components are put on test and that their lifetimes  $X_1, X_2, \dots, X_n$  follow the generalized exponential (GE) distribution with cumulative distribution function (cdf)

$$F(x; \alpha, \lambda) = (1 - e^{-\lambda x})^\alpha$$

and probability density function (pdf)

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

where  $\alpha$  is the shape parameter and  $\lambda$  the reciprocal of a scale parameter. When  $\alpha = 1$ , the GE can be used quite effectively in analyzing any lifetime data, especially in the presence of censoring or if the data is grouped. For distribution (1), the hazard function, which plays an important role in life testing and reliability problems, can be increasing, decreasing or constant depending on the shape parameter  $\alpha$ . For any  $\lambda$ ; the hazard function is nondecreasing if  $\alpha > 1$ ; it is nonincreasing if  $\alpha < 1$  and it is constant if  $\alpha = 1$ . The GE is a unimodal density function and for fixed scale parameter, as the shape parameter increases it becomes more and more symmetric. In addition, the GE distribution has a good physical interpretation. If the lifetime distribution of each component in a parallel system is GE, then the time distribution of the system is also GE. If the data coming from a right-tailed distribution, then the GE can be used quite effectively. It is observed that the GE distribution provides a better fit than the two-parameter gamma and Weibull distributions. Different problems related to GE have been discussed in the literature; see for example, Gupta & Kundu<sup>[6,7,8,9]</sup> and Raqab<sup>[11]</sup>.

This paper is concerned with the problem of estimating the unknown parameters of a mixed (GE) population model based upon a sample censored at a fixed test termination time i.e. type-I censored. Attention will be directed primarily to the case of two subpopulation of failure, each generalized exponentially distributed from subpopulation, element of information matrix obtained, and a numerical illustration will be carried out.

**The Maximum Likelihood Estimators:** In censored type I samples, a random sample of  $n$  units is drawn from the population and placed on test, the test is terminated at fixed time  $T$ , at which time,  $r$  units have failed,  $r_1$  units from subpopulation (1),  $r_2$  units failing from subpopulation (2),  $r_1 + r_2 = r$ , and  $(n-r)$  units which have not failed yield no information concerning the sub-population which they belong to. Following Mendenhall & Hader<sup>[10]</sup>. The likelihood function will be

$$L = \frac{n!}{(n-r)!} [G(T)]^{n-r} p^x q^x \prod_{j=1}^{r_1} f_1(x_{1j}) \prod_{j=1}^{r_2} f_2(x_{2j}),$$

This can be written as

$$L = \frac{n!}{(n-r)!} [1 - \sum_{i=1}^2 p_i F_i(T)]^{n-r} p^x q^x \prod_{i=1}^2 \prod_{j=1}^{r_i} f_i(x_{ij})$$

Using density function (1), the likelihood function will be

$$L = \frac{n!}{(n-r)!} [1 - \sum_{i=1}^2 p_i F_i(T)]^{n-r} p^{r_1} q^{r_2} \prod_{i=1}^2 \prod_{j=1}^{r_i} \alpha_i \lambda_i (1 - e^{-\lambda_i x_{ij}})^{\alpha_i - 1} e^{-\lambda_i x_{ij}}$$

Take the logarithmic

$$\begin{aligned} \ln L = \ln C + (n-r) \ln G + \sum_{i=1}^2 r_i \ln p_i + r_i \ln \alpha_i + r_i \ln \lambda_i \\ + \sum_{i=1}^2 \sum_{j=1}^{r_i} [(\alpha_i - 1)(1 - e^{-\lambda_i x_{ij}}) - \lambda_i x_{ij}] \end{aligned}$$

Where,

$$C = \frac{n!}{(n-r)!} \text{ and } G(T) = 1 - \sum_{i=1}^2 p_i F_i(T) = 1 - \sum_{i=1}^2 p_i (1 - e^{-\lambda_i T})^{\alpha_i}$$

Taking the first partial derivatives of  $\ln L$  with respect to  $p_i, \alpha_i, \lambda_i$ , and  $\lambda_i$ ,

$$\frac{\partial \ln L}{\partial p} = \frac{(n-r)}{G} \frac{\partial G}{\partial p} + \frac{r_1}{p} - \frac{r_2}{q},$$

$$\frac{\partial \ln L}{\partial \alpha_i} = \frac{(n-r)}{G} \frac{\partial G}{\partial \alpha_i} + \frac{r_i}{\alpha_i} + \sum_{j=1}^{r_i} \ln(1 - e^{-\lambda_i x_{ij}}),$$

$$\text{and } \frac{\partial \ln L}{\partial \lambda_i} = \frac{(n-r)}{G} \frac{\partial G}{\partial \lambda_i} + \frac{r_i}{\lambda_i} - r_i \bar{x}_i + (\alpha_i - 1) \sum_{j=1}^{r_i} \frac{x_{ij} e^{-\lambda_i x_{ij}}}{(1 - e^{-\lambda_i x_{ij}})}, \quad i = 1, 2 \quad (2)$$

Where,

$$\frac{\partial G}{\partial p} = (1 - e^{-\lambda_2 T})^{\alpha_2} - (1 - e^{-\lambda_1 T})^{\alpha_1}, \quad \frac{\partial G}{\partial \lambda_i} = \alpha_i T p e^{-\lambda_i T} (1 - e^{-\lambda_i T})^{\alpha_i - 1},$$

$$\text{and } \frac{\partial G}{\partial \alpha_i} = -p_i (1 - e^{-\lambda_i T})^{\alpha_i} \ln(1 - e^{-\lambda_i T}), \quad i = 1, 2$$

Equating equations (2) to zero, and solving to obtain the MLE  $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\lambda}_1, \hat{\lambda}_2$ , and  $\hat{p}$  of the parameters  $\alpha_1, \alpha_2, \lambda_1, \lambda_2$ , and  $p$ , respectively. No explicit solution for this system of equations, numerical method and computer facilities are needed to solve this system of equations.

**Asymptotic Variance-covariance Matrix:** The asymptotic variance-covariance matrix can be obtained by inverting the information matrix with the elements that are negative of the expected values of the second order derivative of logarithms of the likelihood functions. Cohen<sup>[2]</sup> concluded that the approximate variance covariance matrix may be obtained by replacing expected values by their MLE's.

Using the logarithm of the likelihood function, the element of the information matrix are given by

$$\lambda_1, \lambda_2 \frac{\partial^2 \ln L}{\partial p^2} = \frac{-(n-r)}{G^2} \left( \frac{\partial G}{\partial p} \right)^2 - \frac{r_1}{p^2} + \frac{r_2}{q^2}, \quad \frac{\partial^2 \ln L}{\partial p \partial \lambda_i} = \frac{(n-r)}{G} \left( \frac{\partial^2 G}{\partial p \partial \lambda_i} - \left( \frac{\partial G}{\partial p} \right) \left( \frac{\partial G}{\partial \lambda_i} \right) \frac{1}{G} \right),$$

$$\frac{\partial^2 \ln L}{\partial p \partial \alpha_i} = \frac{(n-r)}{G} \left( \frac{\partial^2 G}{\partial p \partial \alpha_i} - \left( \frac{\partial G}{\partial p} \right) \left( \frac{\partial G}{\partial \alpha_i} \right) \frac{1}{G} \right), \quad \frac{\partial^2 \ln L}{\partial \alpha_i^2} = \frac{(n-r)}{G} \left( \frac{\partial^2 G}{\partial \alpha_i^2} - \left( \frac{\partial G}{\partial \alpha_i} \right)^2 \left( \frac{1}{G} \right) \right) - \frac{r_i}{\alpha_i^2}, i=1,2$$

$$2, \frac{\partial^2 \ln L}{\partial \alpha_1 \partial \lambda_2} = \frac{-(n-r)}{G^2} \left( \frac{\partial G}{\partial \alpha_1} \right) \left( \frac{\partial G}{\partial \lambda_2} \right), \quad \frac{\partial^2 \ln L}{\partial \alpha_1 \partial \alpha_2} = \frac{-(n-r)}{G^2} \left( \frac{\partial G}{\partial \alpha_1} \right) \left( \frac{\partial G}{\partial \alpha_2} \right),$$

$$\frac{\partial^2 \ln L}{\partial \lambda_1 \partial \lambda_2} = \frac{-(n-r)}{G^2} \left( \frac{\partial G}{\partial \lambda_2} \right) \left( \frac{\partial G}{\partial \lambda_1} \right), \quad \frac{\partial^2 \ln L}{\partial \lambda_1 \partial \alpha_2} = \frac{-(n-r)}{G^2} \left( \frac{\partial G}{\partial \lambda_1} \right) \left( \frac{\partial G}{\partial \alpha_2} \right)$$

$$\frac{\partial^2 \ln L}{\partial \lambda_i^2} = \frac{(n-r)}{G} \left( \frac{\partial^2 G}{\partial \lambda_i^2} - \left( \frac{\partial G}{\partial \lambda_i} \right)^2 \left( \frac{1}{G} \right) \right) - \frac{r_i}{\lambda_i^2} - (\alpha_i - 1) \sum_{j=1}^{r_i} \frac{x_{ij}^2 e^{-\lambda_i x_{ij}}}{(1 - e^{-\lambda_i x_{ij}})^2}, \quad i=1,2$$

and

$$\frac{\partial^2 \ln L}{\partial \alpha_i \partial \lambda_i} = \frac{(n-r)}{G} \left( \frac{\partial^2 G}{\partial \alpha_i \partial \lambda_i} - \left( \frac{\partial G}{\partial \alpha_i} \right) \left( \frac{\partial G}{\partial \lambda_i} \right) \frac{1}{G} \right) + \sum_{j=1}^{r_i} \frac{x_{ij} e^{-\lambda_i x_{ij}}}{(1 - e^{-\lambda_i x_{ij}})}, \quad i=1,2,$$

$$(-1)^{i+1} \alpha_i^T e^{-\lambda_i^T} (1 - e^{-\lambda_i^T}) \alpha_i^{-1}, \quad \frac{\partial^2 G}{\partial p \partial \alpha_i} = (-1)^i (1 - e^{-\lambda_i^T}) \alpha_i \ln(1 - e^{-\lambda_i^T}),$$

$$\frac{\partial^2 G}{\partial \alpha_i^2} = -p_i (1 - e^{-\lambda_i^T}) \alpha_i [\ln(1 - e^{-\lambda_i^T})]^2, \quad \frac{\partial^2 G}{\partial \lambda_i^2} = \alpha_i^T p_i e^{-\lambda_i^T} (1 - e^{-\lambda_i^T}) \alpha_i^{-2} (\alpha_i e^{-\lambda_i^T} - 1),$$

and  $\frac{\partial^2 G}{\partial \lambda_i \partial \alpha_i} = T p_i e^{-\lambda_i^T} (1 - e^{-\lambda_i^T}) \alpha_i^{-1} (\alpha_i \ln(1 - e^{-\lambda_i^T}) + 1), \quad i=1,2$

Different special cases can be obtained as follows:

- if  $\alpha = 1$  our results reduces to the mixed exponential results obtained by Mendenhall & Hader<sup>[10]</sup>
- if  $n = r$ , the new results reduces to GE results in case of complete sample,
- if  $\alpha_1 = \alpha_2$  our results reduces to results for two mixed GE with common shape parameter.

**Numerical Illustration:** In this section, we present a numerical example to illustrate the maximum likelihood estimators for the unknown parameters, and then variance covariance matrix. To generate random numbers from GE distribution with unknown parameters  $\lambda$  and  $\alpha$  we have

$$F(x, \alpha, \lambda) = (1 - e^{-\lambda x})^\alpha$$

by taking the logarithm , then

$$\ln F(x) = \alpha \ln(1 - e^{-\lambda x}) \quad (3)$$

No explicit solution for the random variable x and iteration procedure is needed to obtain x for given  $\lambda$  and  $\alpha$ .

The left hand side of (3) is distributed uniform (0,1), for fixed  $\lambda$  and uniform random numbers, x may be obtained . Using the Mathcad 2001 package we generate two sets of data from a mixed population when the failure time distributed general exponential in the case of censored type-I where  $T = 0.9$  ,  $n=50$ . The generated data are as follows:

First subpopulation:  $\alpha_1 = 3, \lambda_1 = 2, p = 0.6, r_1 = 17$

0.057 0.116 0.243 0.299 0.339 0.376 0.399 0.409 0.432 0.515 0.558 0.600 0.728 0.742 0.815 0.831  
0.841

Second population:  $\alpha_2 = 4, \lambda_2 = 2.5, p = 0.6, r_2 = 12$

0.146 0.391 0.410 0.516 0.611 0.659 0.719 0.727 0.754 0.816 0.834 0.847

Therefore  $n - r = 12$ .

Using generated data and mathcad program, the solution of the system (1) is obtained, that is,

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \lambda_1 \\ \lambda_2 \\ p \end{pmatrix} = \begin{pmatrix} 2.274 \\ 1.8 \\ 2.395 \\ 0.86 \\ 0.431 \end{pmatrix}$$

The corresponding approximate variance covariance matrix will be

$$\begin{bmatrix} \text{var}(\hat{\alpha}_1) & \text{cov}(\hat{\alpha}_1, \hat{\alpha}_2) & \text{cov}(\hat{\alpha}_1, \hat{\lambda}_1) & \text{cov}(\hat{\alpha}_1, \hat{\lambda}_2) & \text{cov}(\hat{\alpha}_1, \hat{p}) \\ \text{cov}(\hat{\alpha}_2, \hat{\alpha}_1) & \text{var}(\hat{\alpha}_2) & \text{cov}(\hat{\alpha}_2, \hat{\lambda}_1) & \text{cov}(\hat{\alpha}_2, \hat{\lambda}_2) & \text{cov}(\hat{\alpha}_2, \hat{p}) \\ \text{cov}(\hat{\lambda}_1, \hat{\alpha}_1) & \text{cov}(\hat{\lambda}_1, \hat{\alpha}_2) & \text{var}(\hat{\lambda}_1) & \text{cov}(\hat{\lambda}_1, \hat{\lambda}_2) & \text{cov}(\hat{\lambda}_1, \hat{p}) \\ \text{cov}(\hat{\lambda}_2, \hat{\alpha}_1) & \text{cov}(\hat{\lambda}_2, \hat{\alpha}_2) & \text{cov}(\hat{\lambda}_2, \hat{\lambda}_1) & \text{var}(\hat{\lambda}_2) & \text{cov}(\hat{\lambda}_2, \hat{p}) \\ \text{cov}(\hat{p}, \hat{\alpha}_1) & \text{cov}(\hat{p}, \hat{\alpha}_2) & \text{cov}(\hat{p}, \hat{\lambda}_1) & \text{cov}(\hat{p}, \hat{\lambda}_2) & \text{var}(\hat{p}) \end{bmatrix} = \begin{bmatrix} 1.0422 & 0.0074 & 1.1968 & -0.0839 & -0.1342 \\ 0.0074 & 0.5511 & 0.0149 & 0.2416 & -0.0025 \\ 1.1968 & 0.0148 & 1.9861 & -0.1678 & -0.2684 \\ -0.0839 & 0.2416 & -0.1678 & 0.1553 & 0.0286 \\ -0.1342 & -0.0025 & -0.2684 & 0.0286 & 0.0534 \end{bmatrix}$$

We note that the present result can be extended to more than subpopulation very easily, An extensive simulation procedure is needed to study the properties of the new estimators.

**REFERENCES**

1. Acheson, M.A. and E.M. McElwee, 1951. Concerning the reliability of electron tubes. The Sylvania Technologist, 4.	3. Davies, J.A., 1952. Life test acceptance sampling methods. Unpublished paper presented at a summer statistics conference, University of North Carolina.
2. Cohen, A.C., 1965. Maximum likelihood estimation in the Weibull distribution based on complete and censored samples., Tech., 7: 579.	

4. Epstein, B., 1953. "Statistical problem in life testing". proceeding of the seventh Annual Convention, American Society for Quality Control.
5. Gupta, R.D. and D. Kundu, 1999."Generalized exponential distribution" Australian and New Zealand Journal of Statistics, 41(2): 173-188.
6. Gupta, R.D. and D. Kundu, 2003a. "Discriminating between Weibull and generalized exponential distributions " *Comput. Statist. and Data Anal.*, 43: 179-196.
7. Gupta, R.D. and D. Kundu, 2003b. "Closeness of Gamma and generalized exponential distributions " *Comm. Statist.- theory Methods*, 32(4): 705-721.
8. Gupta, R.D. and D. Kundu, 2004. " Discriminating between Gamma and generalized exponential distributions " *J. Statist. Comput. Simulation* 74(2): 107-121.
9. Gupta, R.D. and D. Kundu, 2007. " Generalized exponential distribution: Existing results and some recent developments" *J. Statist. Plann. Inference.*, 137: 3537-3547.
10. Mendenhall, W. and R.J. Hader, 1958." Estimation of parameters of mixed exponentially Distributed Failure Time Distributions From Censored Life Test Data" *Biometrika*, 45: 504.
11. Raqab, M.Z., 2002. "Inference for generalized exponential distribution based on record statistics". *J. of Statist. Plann. Inference.*, 104: 339-350.
12. Raqab, M.Z. and M.T. Madi, 2005. "Bayesian inference for generalized exponential distribution" *J. Statist. Comput. Simulation.*, 75(10): 841-852.
13. Steen, J.R., 1952. " Life Testing of Electronic tubes." Unpublished paper presented at summer statistics conference, University of North Carolina.
14. Wilde, R.D., 1952. "Nature of rate of failure curves". Unpublished paper presented at a summer statistical conference, University of North Carolina.