Journal of Applied Analysis and Computation
Volume 7, Number 1, February 2017, 346–371

Website:http://jaac-online.com/
DOI:10.11948/2017023

STATISTICAL ANALYSIS TECHNIQUE OF TWO-PARAMETER GENERALIZED EXPONENTIAL SUM DISTRIBUTION*

Ronghua Wang [1] and Beiqing Gu[2,†]

Abstract A new life distribution is proposed, known as "two-parameter generalized exponential sum distribution". We study the density function and failure rate function, the average failure rate function, the image features and the numerical characteristics of the mean residual life of the distribution. Several methods of calculating point estimation of parameters are discussed. Through the Monte-Carlo simulation, we compare the precision of the point estimations. In our opinion, the best linear unbiased estimation is the most optimal solution of these methods. At the same time, several methods of calculating parameters of interval estimations are given. We also discuss the precision of interval estimations by Monte-Carlo simulation and use the best linear unbiased estimation and the best linear invariant estimation to construct interval estimations which are better than other estimation method. Finally, several simulation examples and a case of maintaining tanks is used to illustrate the application of the methods presented in this paper.

Keywords Two-parameter generalized exponential sum distribution, three-parameter generalized exponential sum distribution, exponential sum distribution, image feature, numerical characteristics, point estimation, interval estimation, Monte-Carlo simulation.

MSC(2010) 62N05.

1. Introduction

A function has the following form:

$$F(x) = 1 - [1 + (x/\beta)^m]\, e^{-(x/\beta)^m}, \quad x > 0,$$

where $\beta > 0$ is the scale parameter and $m > 0$ is the shape parameter.

Definition 1.1. If the distribution function of a non-negative continuous random variable is $F(x) = 1 - [1 + (x/\beta)^m]\, e^{-(x/\beta)^m}$,$x > 0$, then X obeys "two-parameter generalized exponential sum distribution", denoted as $X \sim \mathrm{GES}(\beta; m)$.

Conspicuously, if $X \sim \mathrm{GES}(\beta; m)$, the density function $f(x)$, the reliability function $\bar{F}(x)$, the failure rate function $\lambda(x)$, the average failure rate function $G(x)$

†the corresponding author. Email address:gubeiqing@suibe.edu.cn(G. bei)
[1]Mathematics and Science College, Shanghai Normal University, Shanghai 200234, China
[2]School of Statistics and Information, Shanghai University of International Business and Economics, Shanghai 201620, China
*This work was partially supported by Innovation Program of Shanghai Municipal Education Commission (14YZ080) and 2013 National Statistical Scientific Research Plan Key Project (2013LZ08).

and the mean residual life function $m(x)$ of X are as follows, respectively.

$$f(x)=\frac{mx^{2m-1}}{\beta^{2m}}e^{-(x/\beta)^m},\ \bar{F}(x)=\left[1+(x/\beta)^m\right]e^{-(x/\beta)^m},\ \lambda(x)=\frac{m}{\beta^m}\frac{x^{2m-1}}{x^m+\beta^m},$$
$$G(x)=-\frac{1}{x}\ln\bar{F}(x)=-\frac{1}{x}\ln\left[1+(x/\beta)^m\right]-(x/\beta)^m,$$
$$m(x)=\frac{1}{\bar{F}(x)}\int_x^{+\infty}\bar{F}(t)\mathrm{d}t,\ x>0.$$

Especially, when $m=1,\beta=1$, $X\sim\mathrm{GES}(1;1)$, the distribution function and the density function of X are $F(x)=1-e^{-x}-xe^{-x}$, $f(x)=xe^{-x}$, $x>0$, then X obeys "Standard generalized exponential sum distribution".

Obviously, if $X\sim\mathrm{GES}(\beta;m)$, let $Y=(X/\beta)^m$, then $Y\sim\mathrm{GES}(1;1)$.

This paper proposes a new life distribution, known as the "two-parameter generalized exponential sum distribution", and its typical characteristics is that the failure rate function takes the shape of "inverted bathtub" for $1/2<m<1$. We study the density function and failure rate function, the average failure rate function, the image features and the numerical characteristics of the mean residual life of the distribution. Through the Monte-Carlo simulation, we compare the precision of the point estimations. In our opinion, the best linear unbiased estimation is the most optimal solution of these methods. At the same time, several methods of calculating parameters of interval estimations are given. We also discuss the precision of interval estimations by Monte-Carlo simulation and use the best linear unbiased estimation and the best linear invariant estimation to construct interval estimations which are better than other methods. Finally, several simulation examples and a case of maintaining tanks are used to illustrate the application of the method in this paper.

2. "Two-parameter generalized exponential sum distribution", "Three-parameter generalized exponential sum distribution" and their application background

First of all, we explain the origin of "two-parameter generalized exponential sum distribution" and "three-parameter generalized exponential sum distribution".

Definition 2.1. Suppose random variables X,Y are independent of each other, and if $X\sim\mathrm{Exp}(\lambda_1)$, $Y\sim\mathrm{Exp}(\lambda_2)$, then the distribution of random variable $Z=X+Y$ is an "exponential sum distribution".

Obviously, the distribution function $F_Z(z)$ and the density function $f_Z(z)$ are

$$F_Z(z)=1+\frac{\lambda_2e^{-\lambda_1z}}{\lambda_1-\lambda_2}-\frac{\lambda_1e^{-\lambda_2z}}{\lambda_1-\lambda_2},\ f_Z(z)=\frac{\lambda_1\lambda_2}{\lambda_1-\lambda_2}(e^{-\lambda_2z}-e^{-\lambda_1z}),\ z\geq0.$$

Particularly, when $\lambda_1=\lambda_2=\lambda$, $F_Z(z)=1-e^{-\lambda z}-\lambda ze^{-\lambda z}$, $f_Z(z)=\lambda^2ze^{-\lambda z}$.

It is noteworthy that "exponential sum distribution" actually corresponds to the life distribution of cold-standby system which is made up of two independent exponential distribution units. But Ding Yong in the literature [2] names the distribution of the above referred to as "exponential difference distribution", the main reason is that its density function contains the difference between two exponential functions. The literature [7] calls it "sub-exponential distribution" and points out that the computer system's I/O service time of the operation usually follows the "sub-exponential distribution". In this paper we think that the definition: exponential sum distribution is more appropriate. The literature [2] also discusses the "exponential sum distribution" of extremum, inflection point, mathematical expectation and variance, etc. The moment estimation of parameters is given, as well as the relationship between the "exponential sum distribution" and exponential distribution.

From the above, we introduce a shape parameter to the "exponential sum distribution" to get another distribution and name it as "three-parameter generalized exponential sum distribution".

Definition 2.2. If the distribution function of a non-negative continuous random variable X obeys "three-parameter generalized exponential sum distribution", denoted as $X \sim \mathrm{GES}(\beta_1, \beta_2; m)$, its distribution function $F(x)$ and density function $f(x)$ have the following forms:

$$\begin{aligned} F(x) &= 1 + \frac{\beta_1^m e^{-(x/\beta_1)^m}}{\beta_2^m - \beta_1^m} - \frac{\beta_2^m e^{-(x/\beta_2)^m}}{\beta_2^m - \beta_1^m}, \\ f(x) &= \frac{m}{\beta_2^m - \beta_1^m} x^{m-1} \left[e^{-(x/\beta_2)^m} - e^{-(x/\beta_1)^m} \right], \end{aligned}$$

where $m > 0$ is the shape parameter and $\beta_1 > 0, \beta_2 > 0$ are the scale parameters.

This distribution is called "three-parameter generalized exponential sum distribution", if $\lambda_1 = 1/\beta_1^m, \lambda_2 = 1/\beta_2^m$, the forms of distribution function $F(x)$ and density function $f(x)$ will become:

$$\begin{aligned} F(x) &= 1 + \frac{\lambda_2}{\lambda_1 - \lambda_2} e^{-\lambda_1 x^m} - \frac{\lambda_1}{\lambda_1 - \lambda_2} e^{-\lambda_2 x^m}, \\ f(x) &= \frac{m\lambda_1\lambda_2}{\lambda_1 - \lambda_2} x^{m-1} \left(e^{-\lambda_2 x^m} - e^{-\lambda_1 x^m} \right). \end{aligned}$$

If $m = 1$, then the distribution is "exponential sum distribution".

Furthermore, it is worth noting that $F(x) = 1 - e^{-(x/\beta_2)^m}$ for $\beta_1 \to 0$, which is two-parameter Weibull distribution with shape-scale parameters. That is to say, "three-parameter generalized exponential sum distribution" can also be regarded as the generalization of two-parameter Weibull distribution.

If $\beta_1 = \beta_2 = \beta$, X obeys "two-parameter generalized exponential sum distribution", $X \sim \mathrm{GES}(\beta; m)$. In fact, if $\lambda_2 = \lambda_1 = \lambda = 1/\beta^m$, then

$$\begin{aligned} F(x) &= \lim_{\lambda_2 \to \lambda_1} \left(1 + \frac{\lambda_2 e^{-\lambda_1 x^m}}{\lambda_1 - \lambda_2} - \frac{\lambda_1 e^{-\lambda_2 x^m}}{\lambda_1 - \lambda_2} \right) \\ &= 1 - e^{-\lambda x^m} - \lambda x^m e^{-\lambda x^m} = 1 - \left[1 + (x/\beta)^m\right] e^{-(x/\beta)^m}, \\ f(x) &= m x^{m-1} \lambda_1 \lim_{\lambda_2 \to \lambda_1} \frac{\lambda_2 \left(e^{-\lambda_2 x^m} - e^{-\lambda_1 x^m} \right)}{\lambda_1 - \lambda_2} \\ &= m\lambda^2 x^{2m-1} e^{-\lambda x^m} = \frac{m x^{2m-1}}{\beta^{2m}} e^{-(x/\beta)^m}. \end{aligned}$$

We can see from the above, if the shape parameter is $m = 1$, the "three-parameter generalized exponential sum distribution" is the "exponential sum distribution" at the moment. Next, we will illustrate the application of "three-parameter generalized exponential sum distribution" in pharmacokinetics by using this special case of "exponential sum distribution".

The classic pharmacokinetic is based on the compartment model. According to the literature [5] and [13], the linear orally or intramuscular injection —the compartment model, the relationship between drug concentration and time in the body can be modeled by:

$$c(t) = \frac{\lambda_a FD}{V(\lambda_a - \lambda)} \left(e^{-\lambda t} - e^{-\lambda_a t}\right),$$

where D stands for dosage of drugs, F is the fraction that is absorbed by the body, V is the apparent volume of distribution of the body, λ_a stands for constant rate of drug absorption, λ is the constant rate of elimination of drugs. Under normal circumstances, drug absorption is faster than elimination, hence $\lambda_a > \lambda$.

To avoid the influence due to selections of compartment numbers on the model, the statistical moment theory is applied to the pharmacokinetic study. The theory tells us that when drugs enter the organisms, the individual differences, biochemical drugs, pharmacology and other random factors affect the length of residence time of each drug molecule in vivo, which can be treated as a random variable and it reflects the drug in the body of the absorption, distribution and elimination of the corresponding overall effect. To convert the medicine-time curve to the probability density curve, $f(t) = \frac{c(t)}{AUC}$ is defined as the probability density of the retention time of drug in the body, where $AUC = \int_0^{+\infty} c(t)\mathrm{d}t$ is the area under the medicine-time curve.

$$\begin{aligned} AUC &= \int_0^{+\infty} c(t)\mathrm{d}t = \int_0^{+\infty} \frac{\lambda_a FD}{V(\lambda_a - \lambda)} \left(e^{-\lambda t} - e^{-\lambda_a t}\right) \mathrm{d}t \\ &= \frac{\lambda_a FD}{V(\lambda_a - \lambda)} \left(\frac{1}{\lambda} - \frac{1}{\lambda_a}\right) = \frac{FD}{V\lambda}. \end{aligned}$$

Then,

$$f(t) = \frac{c(t)}{AUC} = \frac{V\lambda}{FD} \frac{\lambda_a FD}{V(\lambda_a - \lambda)} \left(e^{-\lambda t} - e^{-\lambda_a t}\right) = \frac{\lambda_a \lambda}{\lambda_a - \lambda} \left(e^{-\lambda t} - e^{-\lambda_a t}\right),$$

which is exactly the "exponential sum distribution". The average retention time of drug in the body is its mathematical expectation $\int_0^{+\infty} tf(t)\mathrm{d}t = 1/\lambda_a + 1/\lambda$, which is the sum of the average absorption time $1/\lambda_a$ and the mean time to eliminate drugs $1/\lambda$. The variance of residence time is $1/\lambda_a^2 + 1/\lambda^2$, and the drug concentration reaches their peak (maximum point) is $(\ln \lambda_a - \ln \lambda)/(\lambda_a - \lambda)$, which is exactly the reciprocal of logarithmic average of two rate constants λ_a and λ.

Next, we illustrate its application in the maintenance theory by a particular case of the "two-parameter generalized exponential sum distribution".

For the "two-parameter generalized exponential sum distribution", let $m = 1$, $\beta = \beta_0/2$, that is the Зрланга distribution, and the distribution function and density function are:

$$F(x) = 1 - (1 + 2x/\beta_0)\, e^{-2x/\beta_0}, \ f(x) = (4x/\beta_0^2) e^{-2x/\beta_0}, \ x > 0, \ \beta_0 > 0.$$

Russia introduced the Зрланга distribution when they studied the time of repairing weapons and equipment, and this distribution plays an important role in equipment maintenance theory. The literature [8] analyses the characteristics of Зрланга distribution. Besides, the maximum likelihood method is used to estimate the parameters of the distribution in the full sample situation, and the examples are used to verify the feasibility and practicability of this distribution. The principle of the simulation to the combat damage parts of armored equipment and the computing method of working time for fixing the combat damage parts of armored equipment has been given in literature [11], according to the need of the prediction of maintenance support in wartime, computer simulation is carried out in a certain type of tanks. Besides, the simulation data of repairing tank's time is produced, the time of repairing injured tank is found to obey Зрланга distribution by statistical analysis. In the literature [3], they study the interval estimation of small sample and test problems of Зрланга distribution and point out that the estimated accuracy of the time of equipment maintenance with Зрланга distribution estimation is higher than exponential distribution. Literature [4] gives the maximum likelihood estimation of the parameters under the type-II censoring and investigates the estimation precision through a lot of Monte-Carlo simulations. Secondly, the inverse moment estimations of the parameters are given in the full sample and compared with the moment estimations and maximum likelihood estimations, the moment estimations and maximum likelihood estimations are slightly better than the inverse moment estimations; Finally, the precise interval estimations and the approximate interval estimations of the parameters are acquired. Comparing the accuracy of these two kinds of interval estimations, the precise interval estimation is better than the approximate interval estimation.

3. Some typical characteristics of the "two-parameter generalized exponential sum distribution"

It is easy to obtain Theorem 3.1 and Theorem 3.2 as following.

Theorem 3.1. *Suppose the non-negative continuous random variable $X \sim GES(\beta; m)$, the image of its density function has the following characteristics: (1) when $0 < m \le 1/2$, $f(x)$ is monotone decreasing; (2) when $m > 1/2$, $f(x)$ is increased at first and then decreased.*

Theorem 3.2. *Suppose the non-negative continuous random variable $X \sim GES(\beta; m)$, the image of its failure rate function $\lambda(x)$ has the following characteristics: (1) when $m \ge 1$, $\lambda(x)$ is monotone increasing and belongs to increasing failure rate; (2) when $0 < m \le 1/2$, $\lambda(x)$ is monotone decreasing and belongs to decreasing failure rate; (3) when $1/2 < m < 1$, $\lambda(x)$ is "inverted bathtub" shape with its vertex at $x_0 = \left(\frac{2m-1}{1-m}\right)^{1/m} \beta$.*

The "inverted bathtub" shape of failure rate function is more common in practical problems, such as the "logarithmic normal distribution", which has a wide application in the reliability engineering, and the shape of its failure rate function shows "inverted bathtub". Literature [9] and [6] point out that the "inverted bathtub" shaped of failure rate model can be used to fit the number of survival analysis of organisms (including humans) and so on. Taking into account the bottom of the "inverted bathtub" shape which is the stable stage of the failure rate function, it plays a key role in practical application. Next, the width of the upper bottom of the "inverted bathtub" (referred to as the upper bottom width) is defined as:

Definition 3.1. Suppose $\lambda(x)$ is the failure rate function of the non-negative continuous random variable X, which is of "inverted bathtub" shape, we obtain its maximum value at $x = x_0$. For any given $\varepsilon > 0$, consider the difference $l(\varepsilon) = x_2 - x_1$ between the two roots $x_1, x_2 (x_1 < x_0 < x_2)$ that meets $\lambda(x) - \lambda(x_0) = \varepsilon$, which reflects the width of the upper bottom of "inverted bathtub" under the given precision ε, namely the upper bottom width.

Theorem 3.3. *Suppose the non-negative continuous random variable $X \sim GES(\beta; m)$, when $1/2 < m < 1$, the failure rate function $\lambda(x)$ is shape of "inverted bathtub", for any given $\varepsilon > 0$, the width $l(\varepsilon)$ of the upper bottom of "inverted bathtub" is proportional to the value of the parameter β.*

Proof. Since $\lambda(x)$ has the shape of "inverted bathtub" when $1/2 < m < 1$, for any given $\varepsilon > 0$, there are two roots which satisfy $\lambda(x) - \lambda(x_0) = \varepsilon$, namely $x_1, x_2 (0 < x_1 < x_2)$, and

$$x_0 = \left(\frac{2m-1}{1-m}\right)^{1/m} \beta, \ \lambda(x_0) = \frac{1}{\beta} \frac{(2m-1)^{(2m-1)/m}}{(1-m)^{(m-1)/m}},$$
$$\lambda(x) - \lambda(x_0) = \frac{m}{\beta^m} \frac{x^{2m-1}}{x^m + \beta^m} - \frac{1}{\beta} \frac{(2m-1)^{(2m-1)/m}}{(1-m)^{(m-1)/m}} = \varepsilon,$$
$$\frac{(x/\beta)^{m-1}}{1 + (x/\beta)^{-m}} = \frac{(2m-1)^{(2m-1)/m}}{m(1-m)^{(m-1)/m}} + \frac{\beta}{m}\varepsilon.$$

Denote $t = \frac{x}{\beta}, t_i = \frac{x_i}{\beta}, i = 1, 2$, and the equation about t is : $\frac{t^{m-1}}{1+t^{-m}} = \frac{(2m-1)^{(2m-1)/m}}{m(1-m)^{(m-1)/m}} + \frac{\beta}{m}\varepsilon$. The two roots are t_1, t_2, it is easy to see that when β increases, the right side of the equation also increases, then $t_2 - t_1$ also increases.

So the width $l(\varepsilon) = x_2 - x_1$ of the upper bottom of "inverted bathtub" also increases, then $l(\varepsilon)$ is proportional to the value of the parameter β. □

Theorem 3.4. *Suppose the non-negative continuous random variable $X \sim GES(\beta; m)$, the image of its average failure rate function $G(x)$ has the following characteristics: (1) when $m \geq 1$, $\lambda(x)$ is monotone increasing, and it belongs to increasing failure rate average; (2) when $0 < m \leq 1/2$, $G(x)$ is monotone decreasing and it belongs to decreasing average failure rate; (3) when $1/2 < m < 1$, $G(x)$ is increased at first and then decreased and its shape presents an "inverted bathtub".*

Proof.

$$G(x) = -(\beta^m x)^{-1} \left[\beta^m \ln(x^m + \beta^m) - \beta^m \ln \beta^m - x^m\right],$$
$$G'(x) = \left\{\left[1 + \left(\frac{x}{\beta}\right)^m\right] x^2\right\}^{-1} \times \left\{(m-1)\left(\frac{x}{\beta}\right)^{2m} + \left[1 + \left(\frac{x}{\beta}\right)^m\right] \ln\left[1 + \left(\frac{x}{\beta}\right)^m\right] - \left(\frac{x}{\beta}\right)^m\right\}.$$

Denote $t = (x/\beta)^m > 0$, then $G'(x) = \left[(1+t)x^2\right]^{-1} \left[(m-1)t^2 + (1+t)\ln(1+t) - t\right]$.

Let $g_1(t) = (m-1)t^2 + (1+t)\ln(1+t) - t$, $t > 0$, then

$$g_1'(t) = 2(m-1)t + \ln(1+t), \quad \lim_{t\to 0} g_1(t) = 0,$$

$$\lim_{t\to+\infty} g_1(t) = \lim_{t\to+\infty}\left\{t^2\left[(m-1) + \left(1+\frac{1}{t}\right)\frac{\ln(1+t)}{t} - \frac{1}{t}\right]\right\} = \begin{cases} +\infty, m \geq 1, \\ -\infty, m < 1. \end{cases}$$

Let $g_2(t) = 2(m-1)t + \ln(1+t)$, $t > 0$, then

$$g_2'(t) = 2(m-1) + 1/(1+t), \quad \lim_{t\to 0} g_2(t) = 0,$$

$$\lim_{t\to+\infty} g_2(t) = \lim_{t\to+\infty}\left\{t\left[2(m-1) + \frac{\ln(1+t)}{t}\right]\right\} = \begin{cases} +\infty, m \geq 1, \\ -\infty, m < 1. \end{cases}$$

Let $g_3(t) = 2(m-1) + 1/(1+t)$, $t > 0$, then

$$g_3'(t) = -1/(1+t)^2 < 0, \quad \lim_{t\to 0} g_3(t) = 2m-1, \quad \lim_{t\to+\infty} g_3(t) = 2(m-1).$$

When $m \geq 1$, $g_3(t) > 0$, scilicet $g_2'(t) < 0$, $g_2(t) < 0$, namely $g_1'(t) < 0$, $g_1(t) < 0$, in other words $G'(x) < 0$, that is to say $G(x)$ is monotone increasing and belongs to increasing average failure rate.

When $1/2 < m < 1$, there exists a unique $t_0 = (2m-1)/(2-2m)$, $g_3(t_0) = 0$, $g_3(t) > 0$, $g_2'(t) > 0$, $g_2(t) < 0$, in the $0 < t < t_0$, and $g_3(t) < 0$, $g_2'(t) < 0$ when $t > t_0$, that is to say $g_2(t)$ is increased at first and then decreased, then there exists a unique $t_1(t_1 > t_0 > 0)$ such that $g_2(t_1) = 0$, and $g_2(t) > 0$, $g_1'(t) > 0$ when $0 < t < t_1$, and $g_2(t) < 0$, $g_1'(t) < 0$ when $t > t_1$, it implies $g_1(t)$ is increased at first and then decreased, therefore there also exists a unique $t_2(t_2 > t_1 > t_0 > 0)$such that $g_1(t_2) = 0$, and $g_1(t) > 0$, $G'(x) > 0$ when $0 < t < t_2$, and $g_1(t) < 0$, $G'(x) < 0$, when $t > t_2$, namely $G(x)$ is increased at first and then decreased. □

Theorem 3.5. *Suppose the non-negative continuous random variable $X \sim GES(\beta; m)$, the image of its mean residual life function $m(x)$ has the following characteristics: (1) when $m \geq 1$, $m(x)$ is monotone decreasing; (2) when $0 < m \leq 1/2$, $m(x)$ is monotone increasing; (3) when $1/2 < m < 1$, $m(x)$ is decreased at first and then increased and it takes on a shape of "bathtub".*

Proof.

$$\begin{aligned} m'(x) &= [1-F(x)]^{-2} f(x) \int_x^{+\infty} [1-F(t)]\,\mathrm{d}t - 1 \\ &= \frac{mx^{2m-1}}{\beta^m (x^m+\beta^m)^2 e^{-(x/\beta)^m}} \\ &\quad \times \left[\int_x^{+\infty} (t^m+\beta^m)\, e^{-(t/\beta)^m}\mathrm{d}t - \frac{\beta^m}{m}\frac{(x^m+\beta^m)^2}{x^{2m-1}} e^{-(x/\beta)^m}\right]. \end{aligned}$$

Let $g(x) = \int_x^{+\infty} (t^m+\beta^m)\, e^{-(t/\beta)^m}\mathrm{d}t - \frac{\beta^m}{m}\frac{(x^m+\beta^m)^2}{x^{2m-1}} e^{-(x/\beta)^m}$, $x > 0$.

When $m > 1/2$, we have $\lim_{x\to 0} g(x) = -\infty$ and $\lim_{x\to+\infty} g(x) = 0$.

When $m = 1/2$, we have $\lim_{x\to 0} g(x) = \int_0^{+\infty} (t^m+\beta^m)\, e^{-(t/\beta)^m}\mathrm{d}t - 2\beta^{3/2}$ and $\lim_{x\to+\infty} g(x) = 0$.

When $0 < m < 1/2$, we have $\lim_{x\to 0} g(x) = \int_0^{+\infty} (t^m + \beta^m) e^{-(t/\beta)^m} \mathrm{d}t$ and $\lim_{x\to +\infty} g(x) = 0$.

$$g'(x) = m^{-1} \left[(m-1) + (2m-1)(\beta/x)^m\right] (x^m + \beta^m)(\beta/x)^m e^{-(x/\beta)^m}.$$

Let $h(x) = (m-1)x^m + (2m-1)\beta^m$, $x > 0$.

When $m = 1$, we have $h(x) = (2m-1)\beta^m = \beta^m > 0$.

When $m \neq 1$, we have $h'(x) = m(m-1)x^{m-1}$, $\lim_{x\to 0} h(x) = (2m-1)\beta^m$,

$$\lim_{x\to +\infty} h(x) = \begin{cases} -\infty, \, m < 1, \\ +\infty, \, m > 1. \end{cases}$$

Therefore, when $m \geq 1$, $h(x) > 0$, $g'(x) > 0$, $g(x) < 0$, $m'(x) < 0$, namely $m(x)$ is monotone decreasing; when $1/2 < m < 1$, there exists a unique $x_0 = \left(\frac{2m-1}{1-m}\right)^{1/m} \beta$ such that $h(x_0) = 0$, and in the $0 < x < x_0$, $h(x) > 0$, $g'(x) > 0$, in the $x > x_0$, $h(x) < 0$, $g'(x) < 0$, that is $g(x)$ increases at first and then decreases. Thus, there exists a unique $x_1 (0 < x_1 < x_0)$ such that $g(x_1) = 0$, and in the $0 < x < x_1$, $g(x) < 0$, $m'(x) < 0$, in the $x > x_1$, $g(x) > 0$, $m'(x) > 0$, that is $m(x)$ decreases at first and then increases.

When $0 < m \leq 1/2$, $h(x) < 0$, $g'(x) < 0$, $g(x) > 0$, $m'(x) > 0$, that is to say $m(x)$ is monotone increasing. □

Theorem 3.6. *Suppose the non-negative continuous random variable X obeys the "two-parameter generalized exponential sum distribution" GES$(\beta; m)$, then*

$$E(X^k) = \beta^k \Gamma(2 + k/m).$$

Specifically,

$$D(X) = \beta^2 \left[\Gamma(2 + 2/m) - \Gamma^2(2 + 1/m)\right].$$

4. Parameter estimation of the "two-parameter generalized exponential sum distribution"

4.1. Moment estimation of parameters in the full sample—method one

Let $X_1, X_2, \cdots, X_n$ be a simple random sample of size n from $X \sim \mathrm{GES}(\beta; m)$, the sample observations of them are $x_1, x_2, \cdots, x_n$, the sample mean, the secondary moment and the sample variance are denoted as $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$, $\overline{X^2} = \frac{1}{n}\sum_{i=1}^n X_i^2, S_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$ respectively the observations of them are denoted by $\bar{x}, \overline{x^2}, s_n^2$, respectively.

The following equations can be established by the method of moment estimation:

$$\beta\Gamma(2 + 1/m) = \bar{X}, \beta^2\Gamma(2 + 2/m) = \overline{X^2}.$$

After simplification, the following transcendental equation with the only parameter m is obtained as follows: $\Gamma(2 + 2/m)\Gamma^{-2}(2 + 1/m) = \overline{X^2}/\bar{X}^2$, and the moment estimation of the parameter m is obtained from the solution, which is denoted as $\hat{m}_1$. Then we get the moment estimation of parameter β, denoted as $\hat{\beta}_1 = \bar{X}\Gamma^{-1}(2 + 1/\hat{m}_1)$.

Lemma 4.1. *There is only one positive real root on the transcendental equation* $\Gamma(2+2/m)\Gamma^{-2}(2+1/m) = \overline{X^2}/\bar{X}^2$ *of parameter* m.

Proof. Let $x = 1/m$, then $\Gamma(2+2/m)\Gamma^{-2}(2+1/m) = \Gamma(2+2x)\Gamma^{-2}(2+x)$.

Let $G(x) = \Gamma(2+2x)\Gamma^{-2}(2+x)$, then $\lim_{x\to 0} G(x) = 1$.

Due to

$$\Gamma(x+1) = x\Gamma(x),\ \Gamma(x) = \frac{1}{x}\prod_{i=1}^{+\infty}\left[\left(1+\frac{x}{i}\right)^{-1}\left(1+\frac{1}{i}\right)^{x}\right],$$

then

$$\Gamma(2+2x) = (1+2x)\Gamma(1+2x) = \prod_{i=1}^{+\infty}\left[\left(1+\frac{1+2x}{i}\right)^{-1}\left(1+\frac{1}{i}\right)^{1+2x}\right],$$

$$\Gamma(2+x) = (1+x)\Gamma(1+x) = \prod_{i=1}^{+\infty}\left[\left(1+\frac{1+x}{i}\right)^{-1}\left(1+\frac{1}{i}\right)^{1+x}\right].$$

Furthermore,

$$G(x) = \prod_{i=1}^{+\infty}\left[1+\frac{x^2}{(i+1)^2+2(i+1)x}\right].$$

Since $\frac{x^2}{(i+1)^2+2(i+1)x} > 0$, it is strictly monotone increasing in x, $\sum_{i=1}^{+\infty}\frac{x^2}{(i+1)^2+2(i+1)x}$ is uniform convergent and strictly monotone increasing in x, then $G(x)$ is strictly monotone increasing in x. And as well, $\lim_{x\to+\infty}\frac{x^2}{(i+1)^2+2(i+1)x} = +\infty$, then $\lim_{x\to+\infty} G(x) = +\infty$.

Owing to $\overline{X^2}/\bar{X}^2 > 1$, the equation $\Gamma(2+2/m)\Gamma^{-2}(2+1/m) = \overline{X^2}/\bar{X}^2$ has only one positive real root. □

4.2. The maximum likelihood estimation (MLE) and interval estimation of parameters under the type-II censoring sample—method two

(1) The maximum likelihood estimation of parameters under the type-II censoring sample

Suppose the life of the product $X \sim \mathrm{GES}(\beta; m)$, the product life test will be stopped until the r products out of n products fail, namely the type-II censoring life test, thereby obtaining the first r order failure data $x_{(1)} \le x_{(2)} \le \cdots \le x_{(r)}$, the likelihood function is

$$L(m,\beta) = \frac{n!}{(n-r)!}\frac{m^r}{\beta^{2rm}}\left(\prod_{i=1}^{r} x_{(i)}\right)^{2m-1} \times e^{-\sum_{i=1}^{r}(x_{(i)}/\beta)^m}\left[1+(x_{(r)}/\beta)^m\right]^{n-r} e^{-(n-r)(x_{(r)}/\beta)^m}.$$

Let $\frac{\partial \ln L(m,\beta)}{\partial m}=0$, $\frac{\partial \ln L(m,\beta)}{\partial \beta}=0$, simultaneous equations can be obtained

$$\begin{cases} \frac{r}{m}-2r\ln\beta+2\sum_{i=1}^{r}\ln x_{(i)}-\frac{(n-r)x_{(r)}^{2m}(\ln x_{(r)}-\ln\beta)}{\beta^m\left(\beta^m+x_{(r)}^m\right)} \\ -\sum_{i=1}^{r}\left(\frac{x_{(i)}}{\beta}\right)^m(\ln x_{(i)}-\ln\beta)=0, \\ -\frac{2rm}{\beta}-\frac{(n-r)mx_{(r)}^m}{\beta\left(\beta^m+x_{(r)}^m\right)}+\frac{m}{\beta^{m+1}}\left[\sum_{i=1}^{r}x_{(i)}^m+(n-r)x_{(r)}^m\right]=0. \end{cases}$$

Simplify the second equation, it becomes

$$2r\beta^{2m}+\left(2rx_{(r)}^m-\sum_{i=1}^{r}x_{(i)}^m\right)\beta^m-x_{(r)}^m\left(\sum_{i=1}^{r}x_{(i)}^m+(n-r)x_{(r)}^m\right)=0,$$

$$\Delta=4r\left(2n-r\right)x_{(r)}^{2m}+4rx_{(r)}^m\sum_{i=1}^{r}x_{(i)}^m+\left(\sum_{i=1}^{r}x_{(i)}^m\right)^2>0,$$

$$\beta^m=\frac{1}{4r}\left[-\left(2rx_{(r)}^m-\sum_{i=1}^{r}x_{(i)}^m\right)+\sqrt{4r\left(2n-r\right)x_{(r)}^{2m}+4rx_{(r)}^m\sum_{i=1}^{r}x_{(i)}^m+\left(\sum_{i=1}^{r}x_{(i)}^m\right)^2}\right],$$

then substitute it into the first equation and simplify the following transcendental equation with the only parameter m:$g(m)=0$, where

$$g(m)=r-2r\ln\beta^m+2\sum_{i=1}^{r}\ln x_{(i)}^m-\frac{(n-r)x_{(r)}^{2m}(\ln x_{(r)}^m-\ln\beta^m)}{\beta^m\left(\beta^m+x_{(r)}^m\right)} -\frac{1}{\beta^m}\sum_{i=1}^{r}x_{(i)}^m(\ln x_{(i)}^m-\ln\beta^m),$$

and

$$\beta^m=\frac{1}{4r}\left[-\left(2rx_{(r)}^m-\sum_{i=1}^{r}x_{(i)}^m\right)+\sqrt{4r\left(2n-r\right)x_{(r)}^{2m}+4rx_{(r)}^m\sum_{i=1}^{r}x_{(i)}^m+\left(\sum_{i=1}^{r}x_{(i)}^m\right)^2}\right].$$

Through a lot of Monte-Carlo simulations, the conclusion can be found that the equation $g(m)=0$ has a unique real root. Hence, the root of the transcendental equation $g(m)=0$ of parameter m is the maximum likelihood estimation $\hat{m}_2$, then the maximum likelihood estimation of parameter β is

$$\hat{\beta}_2 = \left\{ \frac{1}{4r} \left[- \left(2rX_{(r)}^{\hat{m}_2} - \sum_{i=1}^{r} X_{(i)}^{\hat{m}_2} \right) + \sqrt{4r(2n-r)X_{(r)}^{2\hat{m}_2} + 4rX_{(r)}^{\hat{m}_2} \sum_{i=1}^{r} X_{(i)}^{\hat{m}_2} + \left(\sum_{i=1}^{r} X_{(i)}^{\hat{m}_2} \right)^2} \right] \right\}^{1/\hat{m}_2}.$$

(2) The maximum likelihood estimation and interval estimation of parameters under the full sample.

In the full sample case, that is to say $r = n$, the likelihood function is

$$L(m, \beta) = m^n \beta^{-2nm} \left(\prod_{i=1}^{n} x_i \right)^{2m-1} e^{-\sum_{i=1}^{n} (x_i/\beta)^m}.$$

Let $\frac{\partial \ln L(m,\beta)}{\partial m} = 0$, $\frac{\partial \ln L(m,\beta)}{\partial \beta} = 0$, then

$$\frac{n}{m} - 2n \ln \beta + 2 \sum_{i=1}^{n} \ln x_i - \sum_{i=1}^{n} \left(\frac{x_i}{\beta} \right)^m (\ln x_i - \ln \beta) = 0,$$

$$-\frac{2nm}{\beta} + \frac{m}{\beta^{m+1}} \sum_{i=1}^{n} x_i^m = 0.$$

We can work out $\beta^m = \frac{1}{2n} \sum_{i=1}^{n} x_i^m$ from the above formula, and the transcendental equation has only the parameter m

$$\frac{\sum_{i=1}^{n} x_i^m \ln x_i}{\sum_{i=1}^{n} x_i^m} - \frac{1}{2m} = \frac{1}{n} \sum_{i=1}^{n} \ln x_i.$$

The maximum likelihood estimation of the parameter m can be obtained from the above equation, denote as $\hat{m}_2$. Besides, the maximum likelihood estimation of the parameterβ can be obtained

$$\hat{\beta}_2 = \left(\frac{1}{2n} \sum_{i=1}^{n} X_i^{\hat{m}_2} \right)^{1/\hat{m}_2}.$$

Lemma 4.2. *The equation of the parameter m:*

$$\frac{\sum_{i=1}^{n} x_i^m \ln x_i}{\sum_{i=1}^{n} x_i^m} - \frac{1}{2m} = \frac{1}{n} \sum_{i=1}^{n} \ln x_i,$$

has only one positive real root.

Proof. Sorting the sample observations $x_1, x_2, \cdots, x_n$ from smallest to largest, denote as $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$, in this case the equation becomes

$$\frac{\sum_{i=1}^{n} x_{(i)}^m \ln x_{(i)}}{\sum_{i=1}^{n} x_{(i)}^m} - \frac{1}{2m} = \frac{1}{n} \sum_{i=1}^{n} \ln x_{(i)}.$$

Let

$$g(m)=\frac{\sum_{i=1}^{n}x_{(i)}^{m}\ln x_{(i)}}{\sum_{i=1}^{n}x_{(i)}^{m}}-\frac{1}{2m},\ m>0,$$

$$\lim_{m\to 0}g(m)=-\infty,\quad \lim_{m\to+\infty}g(m)=\ln x_{(n)}>\frac{1}{n}\sum_{i=1}^{n}\ln x_{(i)},$$

$$g'(m)=\frac{1}{\left(\sum_{i=1}^{n}x_{(i)}^{m}\right)^{2}}\left[\left(\sum_{i=1}^{n}x_{(i)}^{m}\ln^{2}x_{(i)}\right)\left(\sum_{j=1}^{n}x_{(j)}^{m}\right)-\left(\sum_{i=1}^{n}x_{(i)}^{m}\ln x_{(i)}\right)\left(\sum_{j=1}^{n}x_{(j)}^{m}\ln x_{(j)}\right)\right]+\frac{1}{2m^{2}}.$$

Considering

$$\left(\sum_{i=1}^{n}x_{(i)}^{m}\ln^{2}x_{(i)}\right)\left(\sum_{j=1}^{n}x_{(j)}^{m}\right)-\left(\sum_{i=1}^{n}x_{(i)}^{m}\ln x_{(i)}\right)\left(\sum_{j=1}^{n}x_{(j)}^{m}\ln x_{(j)}\right),$$

the terms contains $x_{(i)}^{m}x_{(j)}^{m}$ are

$$\left[\left(\ln^{2}x_{(i)}+\ln^{2}x_{(j)}\right)-2(\ln x_{(i)})(\ln x_{(j)})\right]x_{(i)}^{m}x_{(j)}^{m}=(\ln x_{(i)}-\ln x_{(j)})^{2}x_{(i)}^{m}x_{(j)}^{m}>0.$$

Therefore, $g'(m)>0$, furthermore the equation $g(m)=\frac{1}{n}\sum_{i=1}^{n}\ln x_{(i)}$ has only one positive root. □

Lemma 4.3. *Suppose the non-negative continuous random variable $X\sim GES(\beta;m)$, $X_1,X_2,\cdots,X_n$ is a simple random sample of size n from population X, the sample observations of them are denoted as $x_1,x_2,\cdots,x_n$, $\hat{m}_2$ is the maximum likelihood estimation of the parameter m, then $\hat{m}_2/m,\hat{m}_2(\ln\hat{\beta}_2-\ln\beta)$ are pivotal quantities.*

Proof. $\hat{m}_2$ is the root of the following equation: $\frac{\sum_{i=1}^{n}x_{(i)}^{m}\ln x_{(i)}}{\sum_{i=1}^{n}x_{(i)}^{m}}-\frac{1}{2m}=\frac{1}{n}\sum_{i=1}^{n}\ln x_{(i)}$, that is

$$\frac{\sum_{i=1}^{n}x_{(i)}^{\hat{m}_2}\ln x_{(i)}}{\sum_{i=1}^{n}x_{(i)}^{\hat{m}_2}}-\frac{1}{2\hat{m}_2}=\frac{1}{n}\sum_{i=1}^{n}\ln x_{(i)},$$

Let $Y=(X/\beta)^m, Y_i=(X_i/\beta)^m, i=1,2,\cdots,n$, $Y\sim$ GES$(1;1)$, $Y_1,Y_2,\cdots,Y_n$ are independently and obey GES$(1;1)$, and

$$\frac{\sum_{i=1}^{n}x_{(i)}^{\hat{m}_2}\ln x_{(i)}^{m}}{\sum_{i=1}^{n}x_{(i)}^{\hat{m}_2}}-\frac{1}{2}\frac{m}{\hat{m}_2}=\frac{1}{n}\sum_{i=1}^{n}\ln x_{(i)}^{m},$$

$$\frac{\sum_{i=1}^{n}\left[(x_{(i)}/\beta)^{m}\right]^{\hat{m}_2/m}\ln\left[(x_{(i)}/\beta)^{m}\right]}{\sum_{i=1}^{n}\left[(x_{(i)}/\beta)^{m}\right]^{\hat{m}_2/m}}-\frac{1}{2}\frac{m}{\hat{m}_2}=\frac{1}{n}\sum_{i=1}^{n}\ln\left[(x_{(i)}/\beta)^{m}\right],$$

that is

$$\frac{\sum_{i=1}^{n}y_{(i)}^{\hat{m}_2/m}\ln y_{(i)}}{\sum_{i=1}^{n}y_{(i)}^{\hat{m}_2/m}}-\frac{1}{2}\frac{1}{\hat{m}_2/m}=\frac{1}{n}\sum_{i=1}^{n}\ln y_{(i)}.$$

Thus, $\hat{m}_2/m$ is a pivotal quantity, as well as

$$\hat{\beta}_2^{\hat{m}_2} = \frac{1}{2n}\sum_{i=1}^{n} X_i^{\hat{m}_2},$$
$$(\hat{\beta}_2/\beta)^{\hat{m}_2} = \frac{1}{2n}\sum_{i=1}^{n} [(X_i/\beta)^m]^{\hat{m}_2/m} = \frac{1}{2n}\sum_{i=1}^{n} Y_i^{\hat{m}_2/m},$$

then $\hat{m}_2(\ln\hat{\beta}_2 - \ln\beta)$ is a pivotal quantity. □

Denote a_α, a'_α as the upper α quartiles of the pivotal quantities $\hat{m}_2/m, \hat{m}_2(\ln\hat{\beta}_2 - \ln\beta)$, respectively. We can get the upper α quartiles with different sample size n through 10000 times Monte-Carlo simulations.

If the confidence level $1-\alpha$ is given, the upper $1-\alpha/2, \alpha/2$ quartiles of $\hat{m}_2/m$ are denoted by $a_{1-\alpha/2}, a_{\alpha/2}$, respectively, and the upper $1-\alpha/2, \alpha/2$ quartiles of $\hat{m}_2(\ln\hat{\beta}_2 - \ln\beta)$ are denoted by$a'_{1-\alpha/2}, a'_{\alpha/2}$, respectively, then

$$P\left(a_{1-\alpha/2} \le \hat{m}_2/m \le a_{\alpha/2}\right) = 1-\alpha, P\left(a'_{1-\alpha/2} \le \hat{m}_2(\ln\hat{\beta}_2 - \ln\beta) \le a'_{\alpha/2}\right) = 1-\alpha.$$

Thus the interval estimation of m, β with the confidence level $1-\alpha$ are respectively

$$\left[\hat{m}_2/a_{\alpha/2}, \hat{m}_2/a_{1-\alpha/2}\right], \quad \left[\hat{\beta}_2 e^{-a'_{\alpha/2}/\hat{m}_2}, \hat{\beta}_2 e^{-a'_{1-\alpha/2}/\hat{m}_2}\right].$$

4.3. The inverse moment estimation and interval estimation of parameters under the full sample—method three

Since $X \sim \mathrm{GES}(\beta; m)$, let $Y = (X/\beta)^m$, then $Y \sim \mathrm{GES}(1;1)$, the distribution function and the density function are respectively denoted by

$$F_Y(y) = 1 - e^{-y} - ye^{-y},\ f_Y(y) = ye^{-y}, y > 0,$$
$$E(Y^k) = \Gamma(k+2) = (k+1)!,\ E(Y) = 2,\ E(Y^2) = 6,\ D(Y) = 2.$$

Construct inverse moment estimation equations

$$\frac{1}{n}\sum_{i=1}^{n}(X_i/\beta)^m = 2,\ \frac{1}{n}\sum_{i=1}^{n}(X_i/\beta)^{2m} = 6.$$

Simplify and get the transcendental equation of the parameter m

$$\left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2} = \frac{3}{2n},$$

the inverse moment estimation of the parameter m can be obtained, denote by $\hat{m}_3$.

Besides, we can get the inverse moment estimation of the parameter β

$$\hat{\beta}_3 = \left(\frac{1}{2n}\sum_{i=1}^{n} X_i^{\hat{m}_3}\right)^{1/\hat{m}_3}.$$

Lemma 4.4. *The equation of parameter m:*$\left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2} = \frac{3}{2n}$ *has only one positive real root.*

Proof. Because $X_1, X_2, \cdots, X_n$ are the first r order statistics of a sample size n from the population X, the order statistics of them are denoted by $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$, at this time

$$\left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2} = \frac{3}{2n}$$

can also be denoted by

$$\left(\sum_{i=1}^{n} X_{(i)}^{2m}\right)\left(\sum_{i=1}^{n} X_{(i)}^{m}\right)^{-2} = \frac{3}{2n}.$$

Let

$$g(m) = \left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2}, \ m > 0.$$

It's easy to see $\lim_{m\to 0} g(m) = 1/n$, and if $X_{(n-k+1)} = X_{(n-k+2)} = \cdots = X_{(n)}$, then $\lim_{m\to+\infty} g(m) = 1/k$,

$$g'(m) = 2\left(\sum_{i=1}^{n} X_i^{m}\right)^{-3}\left(\sum_{i=1}^{n} X_i^{2m}\ln X_i \sum_{j=1}^{n} X_j^{m} - \sum_{i=1}^{n} X_i^{2m}\sum_{j=1}^{n} X_j^{m}\ln X_j\right).$$

Notice that in $\sum_{i=1}^{n} X_i^{2m}\ln X_i \sum_{j=1}^{n} X_j^{m} - \sum_{i=1}^{n} X_i^{2m}\sum_{j=1}^{n} X_j^{m}\ln X_j$,

$$\begin{aligned}&X_i^{2m}X_j^{m}\ln X_i + X_j^{2m}X_i^{m}\ln X_j - X_i^{2m}X_j^{m}\ln X_j - X_j^{2m}X_i^{m}\ln X_i\\ =&X_i^{m}X_j^{m}(\ln X_i - \ln X_j)(X_i^{m} - X_j^{m}) = X_i^{m}X_j^{m}(X_i^{m} - X_j^{m})^2\frac{\ln X_i - \ln X_j}{X_i^{m} - X_j^{m}}.\end{aligned}$$

According to Cauchy mean value theorem, there exists ξ, $\min(X_i, X_j) < \xi < \max(X_i, X_j)$, such that $\frac{\ln X_i - \ln X_j}{X_i^m - X_j^m} = \frac{1}{m\xi^m} > 0$, then $g'(m) > 0$ and usually $1/n < 3/(2n) < 1/k$, so the equation has only one positive root. □

Lemma 4.5. *Suppose the non-negative continuous random variable $X \sim GES(\beta; m)$, $X_1, X_2, \cdots, X_n$ is a simple random sample of size n from population X, then $\left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2}$ is a pivotal quantity, which is strictly monotone increasing in m.*

Proof. Let $Y = (X/\beta)^m$, $Y_i = (X_i/\beta)^m$, $i = 1, 2, \cdots, n$, then $Y \sim \text{GES}(1; 1)$, $Y_1, Y_2, \cdots, Y_n$ are independent and obey GES(1; 1).

Due to

$$\begin{aligned}\left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2} &= \left\{\sum_{i=1}^{n}[(X_i/\beta)^m]^2\right\}\left[\sum_{i=1}^{n}(X_i/\beta)^m\right]^{-2}\\ &= \left(\sum_{i=1}^{n} Y_i^{2m}\right)\left(\sum_{i=1}^{n} Y_i^{m}\right)^{-2},\end{aligned}$$

it is easy to see that it is a pivotal quantity and it is strictly monotone increasing in m known by Lemma 4.4. □

The upper α quartile of the pivotal quantity $\left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2}$ is denoted by b_α, then $P\left(\left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2} > b_\alpha\right) = \alpha$, we can get the upper α quartile for different sample size n through 10000 times Monte-Carlo simulation.

The upper $1-\alpha/2,\ \alpha/2$ quartiles of $\left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2}$are denoted by $b_{1-\alpha/2},\ b_{\alpha/2}$, respectively at the confidence level $1-\alpha$,

$$P\left(b_{1-\alpha/2} \le \left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2} \le b_{\alpha/2}\right) = 1-\alpha.$$

Then the interval estimation of parameter m of the confidence level $1-\alpha$ is obtained: $[\hat{m}_{31}, \hat{m}_{32}]$, where $\hat{m}_{31}, \hat{m}_{32}$ are the roots of the $\left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2} = b_{1-\alpha/2}$, $\left(\sum_{i=1}^{n} X_i^{2m}\right)\left(\sum_{i=1}^{n} X_i^{m}\right)^{-2} = b_{\alpha/2}$, respectively.

4.4. The inverse moment estimation and interval estimation of parameters under the type-II censoring sample—method four

Suppose $X_{(1)} \le X_{(2)} \le \cdots \le X_{(r)}$ are the first r order statistic of a sample of size n from the population $X \sim \mathrm{GES}(\beta, m)$, using the method of literature [10] to construct the pivotal quantity of parameter m.

Lemma 4.6. *Suppose $X_{(1)} \le X_{(2)} \le \cdots \le X_{(r)}$ are the first r order statistics of a sample size n from the population $X \sim GES(\beta, m)$, $S_i = \sum_{j=1}^{i} X_{(j)}^m + (n-i)X_{(i)}^m$, $i = 1, 2, \cdots, r$, then the pivotal quantity is denoted by $\sum_{i=1}^{r-1} \ln \frac{S_r}{S_i}$, it is strictly monotone increasing in m, and the transcendental equation $\sum_{i=1}^{r-1} \ln \frac{S_r}{S_i} = c$ of parameter m has a unique positive real root for any arbitrary positive constant $c > 0$.*

Proof. Let $Y = (X/\beta)^m,\ Y_{(i)} = (X_{(i)}/\beta)^m,\ i = 1, 2, \cdots, r,\ Y \sim \mathrm{GES}(1; 1)$, $Y_{(1)}, Y_{(2)}, \cdots, Y_{(r)}$ have the same distribution with the first r order statistics from $X \sim \mathrm{GES}(\beta, m)$ of sample size n.

$$\begin{aligned}\sum_{i=1}^{r-1} \ln \frac{S_r}{S_i} &= \sum_{i=1}^{r-1} \ln \frac{\sum_{j=1}^{r} X_{(j)}^m + (n-r)X_{(r)}^m}{\sum_{j=1}^{i} X_{(j)}^m + (n-i)X_{(i)}^m} \\ &= \sum_{i=1}^{r-1} \ln \frac{\sum_{j=1}^{r} (X_{(j)}/\beta)^m + (n-r)(X_{(r)}/\beta)^m}{\sum_{j=1}^{i} (X_{(j)}/\beta)^m + (n-i)(X_{(i)}/\beta)^m}.\end{aligned}$$

Then, we can easily know $\sum_{i=1}^{r-1} \ln \frac{S_r}{S_i}$ is a pivotal quantity.

Let $X_{(0)} = 0$, because $\sum_{i=1}^{r-1} \ln \frac{S_r}{S_i} = \sum_{i=1}^{r-1} \ln \left[1 + \frac{\sum_{j=i+1}^{r} (n-j+1)\left(X_{(j)}^m - X_{(j-1)}^m\right)}{\sum_{j=1}^{i} (n-j+1)\left(X_{(j)}^m - X_{(j-1)}^m\right)}\right]$.

Let

$$g(m) = \frac{\sum_{j=i+1}^{r} (n-j+1)\left(X_{(j)}^m - X_{(j-1)}^m\right)}{\sum_{j=1}^{i} (n-j+1)\left(X_{(j)}^m - X_{(j-1)}^m\right)}, \lim_{m\to 0} g(m) = 0, \lim_{m\to +\infty} g(m) = +\infty,$$

then

$$\lim_{m\to 0}\sum_{i=1}^{r-1}\ln\frac{S_r}{S_i}=0,$$

$$\lim_{m\to +\infty}\sum_{i=1}^{r-1}\ln\frac{S_r}{S_i}=+\infty, g'(m)=A\left[\sum_{j=1}^{i}(n-j+1)\left(X_{(j)}^m-X_{(j-1)}^m\right)\right]^{-2},$$

where

$$\begin{aligned}A=&\left[\sum_{j=i+1}^{r}(n-j+1)\left(X_{(j)}^m\ln X_{(j)}-X_{(j-1)}^m\ln X_{(j-1)}\right)\right]\\&\times\left[\sum_{k=1}^{i}(n-k+1)\left(X_{(k)}^m-X_{(k-1)}^m\right)\right]-\left[\sum_{j=i+1}^{r}(n-j+1)\left(X_{(j)}^m-X_{(j-1)}^m\right)\right]\\&\times\left[\sum_{k=1}^{i}(n-k+1)\left(X_{(k)}^m\ln X_{(k)}-X_{(k-1)}^m\ln X_{(k-1)}\right)\right].\end{aligned}$$

Note that

$$\begin{aligned}A=&\sum_{j=i+1}^{r}\sum_{k=1}^{i}(n-j+1)(n-k+1)\left(X_{(j)}^m-X_{(j-1)}^m\right)\left(X_{(k)}^m-X_{(k-1)}^m\right)\\&\times\left[\frac{X_{(j)}^m\ln X_{(j)}-X_{(j-1)}^m\ln X_{(j-1)}}{X_{(j)}^m-X_{(j-1)}^m}-\frac{X_{(k)}^m\ln X_{(k)}-X_{(k-1)}^m\ln X_{(k-1)}}{X_{(k)}^m-X_{(k-1)}^m}\right]\\=&\frac{1}{m}\sum_{j=i+1}^{r}\sum_{k=1}^{i}(n-j+1)(n-k+1)\left(X_{(j)}^m-X_{(j-1)}^m\right)\left(X_{(k)}^m-X_{(k-1)}^m\right)\\&\times\left[\frac{X_{(j)}^m\ln X_{(j)}^m-X_{(j-1)}^m\ln X_{(j-1)}^m}{X_{(j)}^m-X_{(j-1)}^m}-\frac{X_{(k)}^m\ln X_{(k)}^m-X_{(k-1)}^m\ln X_{(k-1)}^m}{X_{(k)}^m-X_{(k-1)}^m}\right].\end{aligned}$$

Let $h(x)=x\ln x$, $h'(x)=1+\ln x$. According to Cauchy mean value theorem, there exists a_j, b_k that satisfy $X_{(j-1)}^m<a_j<X_{(j)}^m$, $X_{(k-1)}^m<b_k<X_{(k)}^m$, we have

$$\frac{X_{(j)}^m\ln X_{(j)}^m-X_{(j-1)}^m\ln X_{(j-1)}^m}{X_{(j)}^m-X_{(j-1)}^m}=1+\ln a_j,$$

$$\frac{X_{(k)}^m\ln X_{(k)}^m-X_{(k-1)}^m\ln X_{(k-1)}^m}{X_{(k)}^m-X_{(k-1)}^m}=1+\ln b_k,$$

then

$$\begin{aligned}A=&\frac{1}{m}\sum_{j=i+1}^{r}\sum_{k=1}^{i}(n-j+1)(n-k+1)\left(X_{(j)}^m-X_{(j-1)}^m\right)\\&\times\left(X_{(k)}^m-X_{(k-1)}^m\right)(\ln a_j-\ln b_k)>0.\end{aligned}$$

$g(m)$ is monotone increasing in m and $\sum_{i=1}^{r-1} \ln \frac{S_r}{S_i}$ is also monotone increasing in m.

At the same time it is easy to see, for any $c > 0$, the equation $\sum_{i=1}^{r-1} \ln \frac{S_r}{S_i} = c$ has a unique positive real root. □

Denote the mean value of the pivotal quantity $\sum_{i=1}^{r-1} \ln \frac{S_r}{S_i}$ by c, the inverse moment estimate of parameter m is $\hat{m}_4$, which is the root of the equation

$$\sum_{i=1}^{n-1} \ln \frac{S_n}{S_i} = c.$$

And

$$\frac{S_r}{\beta^m} = \frac{1}{\beta^m}\left[\sum_{j=1}^{r} X_{(j)}^m + (n-r)X_{(r)}^m\right]$$
$$= \sum_{j=1}^{r} (X_{(j)}/\beta)^m + (n-r)(X_{(r)}/\beta)^m = \sum_{j=1}^{r} Y_{(j)} + (n-r)Y_{(r)}$$

is a pivotal quantity, we can get the mean value c' of it by numerical simulation. If now it is in the full sample size case, i.e. $r = n$, at this time $c' = 2n$.

So we can establish the following equation according to the literature [6]: $\frac{1}{\beta^m}\left[\sum_{j=1}^{r} X_{(j)}^m + (n-r)X_{(r)}^m\right] = c'$, (in the full sample,$r = n$,$c' = 2n$), then the inverse moment estimate of the parameter β is $\hat{\beta}_4 = \left\{\frac{1}{c'}\left[\sum_{j=1}^{r} X_{(j)}^{\hat{m}_4} + (n-r)X_{(r)}^{\hat{m}_4}\right]\right\}^{1/\hat{m}_4}$.

Lemma 4.7. *Suppose $X_{(1)} \le X_{(2)} \le \cdots \le X_{(r)}$ are the first r order statistic of a sample size n from the population $X \sim GES(\beta, m)$, $\hat{m}_4/m, \hat{m}_4(\ln\hat{\beta}_4 - \ln\beta)$ are the pivotal quantities.*

Proof. Since $\hat{m}_4$ is the root of the equation $\sum_{i=1}^{r-1} \ln \frac{\sum_{j=1}^{r} X_{(j)}^m + (n-r)X_{(r)}^m}{\sum_{j=1}^{i} X_{(j)}^m + (n-i)X_{(i)}^m} = c$, we have

$$\sum_{i=1}^{r-1} \ln \frac{\sum_{j=1}^{r} X_{(j)}^{\hat{m}_4} + (n-r)X_{(r)}^{\hat{m}_4}}{\sum_{j=1}^{i} X_{(j)}^{\hat{m}_4} + (n-i)X_{(i)}^{\hat{m}_4}} = c,$$
$$\sum_{i=1}^{r-1} \ln \frac{\sum_{j=1}^{r} (X_{(j)}/\beta)^{\hat{m}_4} + (n-r)(X_{(r)}/\beta)^{\hat{m}_4}}{\sum_{j=1}^{i} (X_{(j)}/\beta)^{\hat{m}_4} + (n-i)(X_{(i)}/\beta)^{\hat{m}_4}} = c,$$
$$\sum_{i=1}^{r-1} \ln \frac{\sum_{j=1}^{r} \left[(X_{(j)}/\beta)^m\right]^{\hat{m}_4/m} + (n-r)\left[(X_{(r)}/\beta)^m\right]^{\hat{m}_4/m}}{\sum_{j=1}^{i} \left[(X_{(j)}/\beta)^m\right]^{\hat{m}_4/m} + (n-i)\left[(X_{(i)}/\beta)^m\right]^{\hat{m}_4/m}} = c.$$

Thus, $\hat{m}_4/m$ is a pivotal quantity. And $\frac{1}{\hat{\beta}_4^{\hat{m}_4}}\left[\sum_{j=1}^{r} X_{(j)}^{\hat{m}_4} + (n-r)X_{(r)}^{\hat{m}_4}\right] = c'$.

$$(\beta/\hat{\beta}_4)^{\hat{m}_4}\left\{\sum_{j=1}^{r}\left[(X_{(j)}/\beta)^m\right]^{\hat{m}_4/m} + (n-r)\left[(X_{(r)}/\beta)^m\right]^{\hat{m}_4/m}\right\} = c',$$

then

$$\frac{1}{c'}\left\{\sum_{j=1}^{r}\left[(X_{(j)}/\beta)^m\right]^{\hat{m}_4/m} + (n-r)\left[(X_{(r)}/\beta)^m\right]^{\hat{m}_4/m}\right\} = (\hat{\beta}_4/\beta)^{\hat{m}_4}.$$

We can see that $\hat{m}_4(\ln\hat{\beta}_4 - \ln\beta)$ is a pivotal quantity from the above equation. □

Denote c_α, c'_α as the upper α quartiles of the pivotal quantities $\hat{m}_4/m, \hat{m}_4(\ln\hat{\beta}_4 - \ln\beta)$, respectively, that is $P(\hat{m}_4/m > c_\alpha) = \alpha$, we can get the upper α quartiles for different sample size n and truncating failure number r through 10000 times Monte-Carlo simulation.

If the confidence level $1-\alpha$ is given, the upper $1-\alpha/2, \alpha/2$ quartiles of $\hat{m}_4/m$ are denoted by $c_{1-\alpha/2}, c_{\alpha/2}$, respectively, and the upper $1-\alpha/2, \alpha/2$ quartiles of $\hat{m}_4(\ln\hat{\beta}_4 - \ln\beta)$ are denoted by $c'_{1-\alpha/2}, c'_{\alpha/2}$, respectively, then

$$P\left(c_{1-\alpha/2} \le \hat{m}_4/m \le c_{\alpha/2}\right) = 1-\alpha, P\left(c'_{1-\alpha/2} \le \hat{m}_4(\ln\hat{\beta}_4 - \ln\beta) \le c'_{\alpha/2}\right) = 1-\alpha.$$

The interval estimation of parameters m, β at the confidence level $1-\alpha$ are respectively obtained by: $\left[\hat{m}_4/c_{\alpha/2}, \hat{m}_4/c_{1-\alpha/2}\right]\left[\hat{\beta}_4 e^{-c'_{\alpha/2}/\hat{m}_4}, \hat{\beta}_4 e^{-c'_{1-\alpha/2}/\hat{m}_4}\right]$.

4.5. The best linear unbiased estimation(BLUE) and interval estimation of parameters under the type-II censoring sample—method five

Suppose $X_{(1)} \le X_{(2)} \le \cdots \le X_{(r)}$ are the first r order statistics of a sample size n from the population $X \sim \text{GES}(\beta, m)$, let $Z = \ln X$, $\mu = \ln\beta$, $\sigma = 1/m$, then

$$\begin{aligned} F_Z(z) &= P(\ln X \le z) = P(X \le e^z) = 1 - \left[1 + (e^z/\beta)^m\right] e^{-(e^z/\beta)^m} \\ &= 1 - \left[1 + e^{(z-\mu)/\sigma}\right] e^{-e^{(z-\mu)/\sigma}}. \end{aligned}$$

Therefore, Z is the distribution of location-scale parameter. Thus, according to Gauss-Markov theorem, the best linear unbiased estimations (BLUE) of parameters σ, μ are

$$\begin{aligned} \hat{\sigma} &= \sum_{j=1}^{r} C(n,r,j) Z_{(j)} = \sum_{j=1}^{r} C(n,r,j) \ln X_{(j)}, \\ \hat{\mu} &= \sum_{j=1}^{r} D(n,r,j) Z_{(j)} = \sum_{j=1}^{r} D(n,r,j) \ln X_{(j)}. \end{aligned}$$

The coefficients satisfy

$$\sum_{j=1}^{r} C(n,r,j) = 0, \sum_{j=1}^{r} C(n,r,j)\alpha_j = 1, \sum_{j=1}^{r} D(n,r,j) = 1, \sum_{j=1}^{r} D(n,r,j)\alpha_j = 0,$$

where $C(n,r,j)$ is the coefficient of the best linear unbiased estimation of σ, and $D(n,r,j)$ is the coefficient of the best linear unbiased estimation of μ. We can get the values of the two coefficients in different sample size n and truncating failure number r through 10000 times Monte-Carlo simulations.

Then the estimations of parameters β, m are

$$\hat{\beta}_5 = e^{\hat{\mu}} = \exp\left[\sum_{j=1}^{r} D(n,r,j) \ln X_{(j)}\right], \hat{m}_5 = (\hat{\sigma})^{-1} = \left[\sum_{j=1}^{r} C(n,r,j) \ln X_{(j)}\right]^{-1}.$$

Particularly, in the full sample size $r = n$, $C(n, r, j)$ is denoted by $C(n, j)$ and $D(n, r, j)$ is denoted by $D(n, j)$, then the best linear unbiased estimator (BLUE) of the parameters β, m are

$$\hat{\beta}_5 = e^{\hat{\mu}} = \exp\left[\sum_{j=1}^{n} D(n,j)\ln X_{(j)}\right], \hat{m}_5 = (\hat{\sigma})^{-1} = \left[\sum_{j=1}^{n} C(n,j)\ln X_{(j)}\right]^{-1}.$$

Lemma 4.8. *Suppose $X_{(1)} \le X_{(2)} \le \cdots \le X_{(r)}$ are the first r order statistic of a sample size n from the population $X \sim GES(\beta, m)$, $\hat{\sigma}/\sigma, (\hat{\mu} - \mu)/\hat{\sigma}$ are the pivotal quantities.*

The upper α quartiles of the pivotal quantities $\hat{\sigma}/\sigma, (\hat{\mu} - \mu)/\hat{\sigma}$ are denoted by d_α, d'_α, then $P\left((\hat{\mu} - \mu)/\hat{\sigma} > d'_\alpha\right) = \alpha$. We can get the upper α quartiles for different sample size n and truncating failure number r through 10000 times Monte-Carlo simulation.

If the confidence level $1 - \alpha$ is given, the upper $1 - \alpha/2, \alpha/2$ quartiles of $\hat{\sigma}/\sigma$ are denoted by $d_{1-\alpha/2}, d_{\alpha/2}$, respectively, and the upper $1 - \alpha/2, \alpha/2$ quartiles of $(\hat{\mu} - \mu)/\hat{\sigma}$ are denoted by $d'_{1-\alpha/2}, d'_{\alpha/2}$, respectively, then

$$P\left(d_{1-\alpha/2} \le \hat{\sigma}/\sigma \le d_{\alpha/2}\right) = 1 - \alpha,\ P\left(d'_{1-\alpha/2} \le (\hat{\mu} - \mu)/\hat{\sigma} \le d'_{\alpha/2}\right) = 1 - \alpha.$$

The interval estimations of parameters m, β at the confidence level $1 - \alpha$ are obtained by

$$\left[d_{1-\alpha/2}/\hat{\sigma}, d_{\alpha/2}/\hat{\sigma}\right], \left[\exp\left(\hat{\mu} - d'_{\alpha/2}\hat{\sigma}\right), \exp\left(\hat{\mu} - d'_{1-\alpha/2}\hat{\sigma}\right)\right].$$

4.6. The best linear invariant estimation (BLIE) and the interval estimation of parameters under the type-II censoring sample—method six

It is easy to see that the best linear invariant estimation (BLIE) of parameters σ, μ are

$$\tilde{\sigma} = \sum_{j=1}^{r} C_I(n,r,j)Z_{(j)} = \sum_{j=1}^{r} C_I(n,r,j)\ln X_{(j)},$$
$$\tilde{\mu} = \sum_{j=1}^{r} D_I(n,r,j)Z_{(j)} = \sum_{j=1}^{r} D_I(n,r,j)\ln X_{(j)}.$$

The coefficients of the best linear invariant estimation of the parameters σ, μ are

$$C_I(n,r,j) = C(n,r,j)/(1 + l_{rn}), D_I(n,r,j) = D(n,r,j) - C(n,r,j)B_{rn}/(1 + l_{rn}).$$

Then, the estimations of the parameters β, m are

$$\hat{\beta}_6 = e^{\tilde{\mu}} = \exp\left[\sum_{j=1}^{r} D_I(n,r,j)\ln X_{(j)}\right], \hat{m}_6 = (\tilde{\sigma})^{-1} = \left[\sum_{j=1}^{r} C_I(n,r,j)\ln X_{(j)}\right]^{-1}.$$

In particular, in the full sample size $r = n$,$C_I(n,r,j)$ is denoted by $C_I(n,j)$ and $D_I(n,r,j)$ is denoted by $D_I(n,j)$, then the invariant estimations of the parameters β, m are

$$\hat{\beta}_6 = e^{\tilde{\mu}} = \exp\left[\sum_{j=1}^{n} D_I(n,j)\ln X_{(j)}\right], \hat{m}_6 = (\tilde{\sigma})^{-1} = \left[\sum_{j=1}^{n} C_I(n,j)\ln X_{(j)}\right]^{-1},$$

where $C_I(n,j) = C(n,j)/(1+l_{nn})$, $D_I(n,j) = D(n,j) - C(n,j)B_{nn}/(1+l_{nn})$.

Lemma 4.9. *Suppose $X_{(1)} \le X_{(2)} \le \cdots \le X_{(r)}$ are the first r order statistics of a sample size n from the population $X \sim GES(\beta, m)$, $\tilde{\sigma}/\sigma, (\tilde{\mu}-\mu)/\tilde{\sigma}$ are the pivotal quantities.*

The upper α quartiles of the pivotal quantity $\tilde{\sigma}/\sigma, (\tilde{\mu}-\mu)/\tilde{\sigma}$ are denoted by e_α, e'_α, respectively. We can obtain the upper α quartiles for different sample size n and truncating failure number r though 10000 times Monte-Carlo simulation.

If the confidence level $1-\alpha$ is given, the upper $1-\alpha/2, \alpha/2$ quartiles of $\tilde{\sigma}/\sigma$ are denoted by $e_{1-\alpha/2}, e_{\alpha/2}$, respectively, and the upper $1-\alpha/2, \alpha/2$ quartiles of $(\tilde{\mu}-\mu)/\tilde{\sigma}$ are denoted by $e'_{1-\alpha/2}, e'_{\alpha/2}$, respectively, then

$$P\left(e_{1-\alpha/2} \le \tilde{\sigma}/\sigma \le e_{\alpha/2}\right) = 1-\alpha, P\left(e'_{1-\alpha/2} \le (\tilde{\mu}-\mu)/\tilde{\sigma} \le e'_{\alpha/2}\right) = 1-\alpha.$$

The interval estimations of parameters m, β at the confidence level $1-\alpha$ are obtained by:

$$\left[e_{1-\alpha/2}/\tilde{\sigma}, e_{\alpha/2}/\tilde{\sigma}\right], \left[\exp\left(\tilde{\mu} - e'_{\alpha/2}\tilde{\sigma}\right), \exp\left(\tilde{\mu} - e'_{1-\alpha/2}\tilde{\sigma}\right)\right].$$

4.7. The approximate maximum likelihood estimation (ALME) of the parameter under the type-II censoring sample—method seven

In the literature [1], it points out that, for some distribution, such as exponential distribution, Rayleigh distribution and Weibull distribution, their maximum likelihood estimation of the parameter does not have an explicit expression. Therefore, in order to improve the method of maximum likelihood estimation, the approximate maximum likelihood estimation method is proposed. Since the maximum likelihood estimations of the parameters β, m are the solution of a transcendental equation without explicit expression, according to literature [1], we can get the approximate maximum likelihood estimation of the parameters

$$\hat{\sigma} = \frac{-D + \sqrt{D^2 + 4AE}}{2A}, \ \hat{\mu} = B - C\hat{\sigma},$$

where

$$B = \frac{1}{M}\left[\sum_{i=1}^{r} \eta_i \ln X_{(i)} + (n-r)\eta \ln X_{(r)}\right],$$

$$C = \frac{1}{M}\left[\sum_{i=1}^{r} \gamma_i - (n-r)\gamma\right],$$

$$M=\sum_{i=1}^{r}\eta_i+(n-r)\eta, A=r+C\sum_{i=1}^{r}(\gamma_i-\eta_iC)-(n-r)C(\gamma+\eta C),$$
$$D=\sum_{i=1}^{r}(2\eta_iC-\gamma_i)\left(B-\ln X_{(i)}\right)+(n-r)(2\eta C+\gamma)\left(B-\ln X_{(r)}\right),$$
$$E=\sum_{i=1}^{r}\eta_i\left(B-\ln X_{(i)}\right)^2+(n-r)\eta\left(B-\ln X_{(r)}\right)^2.$$

Since $F_T(t)=1-(1+e^t)\,e^{-e^t}$, $f_T(t)=e^{2t}e^{-e^t}-\infty<t<+\infty$, let $p_i=i/(n+1)$, $q_i=1-p_i$, $i=1,2,\cdots,r$, and ξ_i satisfies $F_T(\xi_i)=p_i$, that is $\left(1+e^{\xi_i}\right)e^{-e^{\xi_i}}=q_i$.

$$\gamma_i=2-e^{\xi_i}+\xi_ie^{\xi_i},\ \eta_i=e^{\xi_i},$$
$$\gamma=\frac{f_T(\xi_r)}{q_r}-\xi_r\frac{f'_T(\xi_r)q_r+[f_T(\xi_r)]^2}{q_r^2},\ \eta=\frac{f'_T(\xi_r)q_r+[f_T(\xi_r)]^2}{q_r^2}.$$

Moreover, the approximate maximum likelihood estimation of the parameters β, m are

$$\hat{\beta}_7=e^{\hat{\mu}}=\exp\left[B-\frac{C\left(\sqrt{D^2+4AE}-D\right)}{2A}\right],\hat{m}_7=\frac{1}{\hat{\sigma}}=\frac{2A}{\sqrt{D^2+4AE}-D}.$$

4.8. Simulation comparison of point estimation of parameter

In the full sample, the accuracy of point estimation of parameter is investigated by Monte-Carlo simulations. Taking the true value of parameters $\beta=1$, $m=1$, 1000 times Monte-Carlo simulations are carried out under different sample sizes. The estimated mean and mean square error of several point estimation methods are calculated, and the simulation results are shown in Table 1 and Table 2. We can see: (1) when the sample size increases, the mean square error of the various point estimation method gradually decreases; (2) for all point estimations of parameter m, the mean square error of BLUE is relatively smaller, and mean square error of each point estimation method is gradually approaching with the increase of sample size; (3) for all point estimations of parameter β, the mean square error of BLUE is relatively smaller, and mean square error of each point estimation method is gradually approaching with the increase of sample size.

4.9. Simulation comparison of interval estimation of parameter

In the full sample, the accuracy of interval estimation of parameter is investigated by Monte-Carlo simulations. Taking the true value of parameters $\beta=1$, $m=1$, 1000 times Monte-Carlo simulations are carried out under different sample sizes. At the confidence level of 0.95, we calculate the average lower limit, the average upper limit, the average interval length and the number of the intervals containing the true value of parameter. The simulation results are shown in Table 3 and Table 4. It can be found that it is better to use the best linear unbiased estimation and the best linear invariant estimation to construct interval estimation when the sample size is smaller from the view of the average interval length. That is to say, the method five

Table 1. Simulation results for point estimation of parameter m

sample size n		10	15	20	25	30
moment estimation—method one	mean	1.1934	1.1322	1.0788	1.0786	1.0607
	mean square error	0.1352	0.0750	0.0474	0.0377	0.0261
maximum likelihood estimation—method two	mean	1.1544	1.0967	1.0692	1.0516	1.0479
	mean square error	0.1294	0.0578	0.0441	0.0272	0.0260
inverse moment estimation—method three	mean	1.2533	1.1637	1.0994	1.0933	1.0716
	mean square error	0.2108	0.1090	0.0632	0.0503	0.0317
inverse moment estimation—method four	mean	1.0889	1.0428	1.0419	1.0319	1.0189
	mean square error	0.1119	0.0499	0.0375	0.0302	0.0213
BLUE—method five	mean	1.0767	1.0376	1.0368	1.0172	1.0168
	mean square error	0.0951	0.0479	0.0367	0.0276	0.0195
BLIE—method six	mean	1.1449	1.0791	1.0671	1.0406	1.0362
	mean square error	0.1218	0.0566	0.0419	0.0302	0.0213
approximate maximum likelihood estimation—method seven	mean	1.1551	1.0832	1.0722	1.0437	1.0399
	mean square error	0.1279	0.0579	0.0429	0.0310	0.0222

Table 2. Simulation results for point estimation of parameter β

sample size n		10	15	20	25	30
moment estimation—method one	mean	1.1032	1.0799	1.0440	1.0511	1.0358
	mean square error	0.1095	0.0739	0.0537	0.0431	0.0338
maximum likelihood estimation—method two	mean	1.0917	1.0499	1.0415	1.0281	1.0283
	mean square error	0.1124	0.0663	0.0519	0.0408	0.0343
inverse moment estimation—method three	mean	1.1395	1.0997	1.0562	1.0604	1.0444
	mean square error	0.1352	0.0904	0.0631	0.0485	0.0370
inverse moment estimation—method four	mean	1.0282	1.0089	1.0132	1.0107	1.0025
	mean square error	0.1091	0.0646	0.0485	0.0428	0.0307
BLUE—method five	mean	1.0488	1.0275	1.0263	1.0095	1.0102
	mean square error	0.1011	0.0645	0.0482	0.0387	0.0299
BLIE—method six	mean	1.1003	1.0611	1.0517	1.0303	1.0273
	mean square error	0.1132	0.0705	0.0512	0.0402	0.0310
approximate maximum likelihood estimation— method seven	mean	1.0529	1.0246	1.0269	1.0090	1.0100
	mean square error	0.1035	0.0652	0.0484	0.0391	0.0304

and method six of interval estimations of parameters m and β are better. Besides, the average interval length of each interval estimation method is gradually close as the sample size increases.

4.10. Fitting test of two-parameter generalized exponential sum distribution under the full sample situation

Suppose non-negative continuous random variable $X \sim \text{GES}(\beta; m)$, $X_1, X_2, \cdots, X_n$ are simple random samples from the population X with sample size n, the order statistics of them are denoted by $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$, and $\hat{m}_2$ is the maximum likelihood estimation of the parameter m. Let $T_i = \sum_{j=1}^{i} X_{(j)}^{\hat{m}_2} + (n-i)X_{(i)}^{\hat{m}_2}, i = 1, 2, \cdots, n$, the distribution of $\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i}$ has nothing to do with the parameter. In fact, let $Y = (X/\beta)^m, Y_{(i)} = (X_{(i)}/\beta)^m, i = 1, 2, \cdots, n$, then $Y \sim \text{GES}(1; 1)$, $Y_{(1)}, Y_{(2)}, \cdots, Y_{(n)}$ are the first$r$order statistics of a sample size n from the popula-

Table 3. Simulation results for interval estimation of parameter m

	n	10	15	20	25	30
interval estimation of method two	average lower limit	0.6018	0.6627	0.7098	0.7344	0.7566
	average upper limit	1.6314	1.4829	1.3945	1.3413	1.2922
	average interval length	1.0296	0.8203	0.6847	0.607	0.5356
	the number of intervals containing the true value	951	965	959	948	957
interval estimation of method three	average lower limit	0.5826	0.6648	0.7084	0.7233	0.7472
	average upper limit	1.8982	1.6617	1.5315	1.4497	1.3992
	average interval length	1.3156	0.9969	0.8231	0.7264	0.6519
	the number of intervals containing the true value	956	947	959	965	957
interval estimation of method four	average lower limit	0.5968	0.6694	0.7078	0.7243	0.7533
	average upper limit	1.6130	1.4872	1.4050	1.3348	1.2960
	average interval length	1.0263	0.8177	0.6972	0.6105	0.5427
	the number of intervals containing the true value	956	950	962	958	956
interval estimation of method five	average lower limit	0.5926	0.6707	0.7066	0.7296	0.7573
	average upper limit	1.6185	1.4875	1.3927	1.3393	1.2932
	average interval length	1.0259	0.8169	0.6862	0.6097	0.5359
	the number of intervals containing the true value	949	950	958	967	960
interval estimation of method six	average lower limit	0.5926	0.6707	0.7064	0.7296	0.7573
	average upper limit	1.6186	1.4875	1.3928	1.3392	1.2933
	average interval length	1.0260	0.8168	0.6864	0.6096	0.5361
	the number of intervals containing the true value	951	950	958	967	960

Table 4. Simulation results for interval estimation of parameter β

	n	10	15	20	25	30
interval estimation of method two	average lower limit	0.4717	0.5513	0.6275	0.6505	0.6697
	average upper limit	1.6947	1.5577	1.4651	1.4341	1.3828
	average interval length	1.223	1.0065	0.8376	0.7837	0.7131
	the number of intervals containing the true value	953	946	954	958	963
interval estimation of method four	average lower limit	0.4602	0.5592	0.6243	0.6478	0.6816
	average upper limit	1.7049	1.5436	1.4595	1.4471	1.3818
	average interval length	1.2447	0.9845	0.8352	0.7993	0.7002
	the number of intervals containing the true value	956	942	952	957	957
interval estimation of method five	average lower limit	0.4890	0.5547	0.6237	0.6480	0.6613
	average upper limit	1.6907	1.5369	1.4665	1.4324	1.3718
	average interval length	1.2018	0.9823	0.8428	0.7843	0.7105
	the number of intervals containing the true value	947	955	951	952	955
interval estimation of method six	average lower limit	0.4889	0.5548	0.6238	0.6480	0.6613
	average upper limit	1.6908	1.5368	1.4662	1.4323	1.3719
	average interval length	1.2020	0.9820	0.8424	0.7843	0.7106
	the number of intervals containing the true value	947	955	951	952	955

tion $Y \sim \mathrm{GES}(1;1)$ with the same distribution.

$$\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i} = \sum_{i=1}^{n-1} \ln \frac{\sum_{j=1}^{n} X_{(j)}^{\hat{m}_2}}{\sum_{j=1}^{i} X_{(j)}^{\hat{m}_2} + (n-i) X_{(i)}^{\hat{m}_2}}$$

$$= \sum_{i=1}^{n-1} \ln \frac{\sum_{j=1}^{r} (X_{(j)}/\beta)^{\hat{m}_2}}{\sum_{j=1}^{i} (X_{(j)}/\beta)^{\hat{m}_2} + (n-i)(X_{(i)}/\beta)^{\hat{m}_2}}$$
$$= \sum_{i=1}^{n-1} \ln \frac{\sum_{j=1}^{r} \left(Y_{(j)}\right)^{\hat{m}_2/m}}{\sum_{j=1}^{i} \left(Y_{(j)}\right)^{\hat{m}_2/m} + (n-i)\left(Y_{(i)}\right)^{\hat{m}_2/m}}.$$

Because $\hat{m}_2/m$ is the pivotal quantity, the distribution of $\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i}$ has nothing to do with the parameter.

Thus, we can get the upper α quartile of the distribution of $\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i}$ for different sample size n through 10000 times Monte-Carlo simulation.

If the confidence level $1-\alpha$ is given, the upper $1-\alpha/2, \alpha/2$ quartiles of $\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i}$ are denoted by $g_{1-\alpha/2}, g_{\alpha/2}$, respectively. Thus, if the observed value of $\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i}$ falls between $(g_{1-\alpha/2}, g_{\alpha/2})$, we can consider that $X_1, X_2, \cdots, X_n$ are simple random samples from the population$X \sim \text{GES}(\beta; m)$ with sample size n.

4.11. Case analysis—the estimation of parameter for the distribution regularity of maintenance time of a certain type tank

The literature [8] investigates a type of tank in the process of maintenance, after the 47 observations of the basic level I preventive maintenance of class two maintenance time, and the field observations are (unit: hours)

0.80	1.00	1.00	1.41	1.50	1.50	1.50	2.00	2.00	2.00
2.00	2.50	2.50	2.75	3.20	3.30	3.70	3.80	3.80	4.00
4.00	4.00	4.00	4.00	4.00	4.10	5.00	5.00	5.50	5.50
5.50	6.00	6.50	7.00	7.16	7.75	8.00	8.00	9.50	9.73
10.00	11.40	12.00	12.00	14.00	15.21	15.50			

For "two-parameter generalized exponential sum distribution", let $m = 1$, $\beta = \beta_0/2$, that is Зрланга distribution, and the distribution function and density function are

$$F(x) = 1 - (1 + 2x/\beta_0)\, e^{-2x/\beta_0},\ \ f(x) = (4x/\beta_0^2)e^{-2x/\beta_0},\ \ x > 0, \beta_0 > 0.$$

The literature [8] considers that the associated fault repair maintenance time of a certain type of tank obeys Зрланга distribution by χ^2 test, and the maximum likelihood estimation of parameter β_0 is $\hat{\beta}_0 = 12.1$. In the literature [4], the inverse moment estimation of parameter β_0 is given by 5.469, and the precise interval estimation is $[4.5104, 6.762]$ and the approximate interval estimation is $[4.5416, 6.8432]$ at the confidence level 95%.

Then we carry out the fitting test of generalized exponential sum distribution. If the confidence level $1 - \alpha = 0.95$ is given, the upper $1 - \alpha/2, \alpha/2$ quartiles of $\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i}$ are denoted by $g_{1-\alpha/2} = 26.4062, g_{\alpha/2} = 28.5391$, respectively, and the observed value of $\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i}$ is 27.7481, which falls between $(g_{1-\alpha/2}, g_{\alpha/2})$, therefore we think that $X_1, X_2, \cdots, X_n$ are simple random samples from $X \sim \text{GES}(\beta; m)$ with sample size n.

Parameter estimation and interval estimation are as follows.

(1) We can get each point estimation of parameters m, β by using the method presented in this paper, as shown in Table 5.

Table 5. The point estimations of parameters m, β

	method one	method two	method three	method four	method five	method six	method seven
point estimation of m	0.9999	1.0051	0.9998	0.9916	0.9820	0.9944	1.0049
point estimation of β	2.7295	2.7427	2.7295	2.7087	2.6805	2.7130	2.7298

From the above point estimates of the parameters, we can find that the point estimation of parameter m is close to 1 and the point estimation of parameter β is close to$\hat{\beta}_0/2 = 2.7297$. From another point of view, it is illustrates that it is reliable to fit the above data by using the GES$(\beta; m)$ distribution.

(2) At a confidence level of 95%, we can get each interval estimation of parameters m, β by using the method presented in this paper, as shown in Table 6.

Table 6. The interval estimations of parameters m, β

		method two	method three	method four	method five	method six
m	interval estimation	0.7734 1.2136	0.7546 1.2594	0.7729 1.2091	0.7677 1.2085	0.7677 1.2085
	length	0.4401	0.5048	0.4362	0.4408	0.4408
β	interval estimation	1.9686 3.5257	—	1.9360 3.5277	1.9116 3.4906	1.9113 3.4899
	length	1.5571	—	1.5917	1.5790	1.5786

References

[1] N. Balakrishnan and A. Clifford Cohen, *Order Statistics and Inference Estimation Methods,* Boston, Academic Press, 1991.

[2] Y. Ding, *Continuous Exponential Difference Distribution*, Journal of Applied of Statistics and Management, 25(2006)(3), 346–350.

[3] G. T. Fan, B. H. Wang and C. L. Chen, *The Research of Interval Estimation and Hypothetical Test of Small Sample of* Зрланга *Distribution,* Journal of Applied of Statistics and Management, 28(2009)(3), 468–472.

[4] B. Q. Gu, R. Wang and X. Xu, *The Statistical Analysis of Distribution,* Conference Proceeding of 2011 National Machinery Industry Reliability Technology Academic Communication, 2011, 65–67.

[5] M. Gibaldi and D. Perrier, *Pharmacokinetic, Second edition*, New York, Marcel Dekker, 1982.

[6] R. C. Gupta and S. Lvin, *Monotonicity of failure rate and mean residual life function of a gamma-type model*, Applied Mathematics and Computation, 165(2005)(3), 623–633.

[7] S. T. Kishor, *Probability and Statistics with Reliability, Queuing, and Computer Science Applications(Second Edition)*, Electronic Industry Press, 188(2015), 110–112.

[8] H. Q. Lv, L. H. Gao and C. L. Chen, Зрланга *Distribution and its Application in Supportability Data Analysis,* Journal of Academy of Armored Force Engineering, 16(2000)(3), 48–52.

[9] J. Wang, H. Muller and W. B. Capre, *Analysis of oldest-old mortality: lifetimes revisited,* Ann. Statist., 26(1998)(1), 126–133.

[10] B. Wang, *Statistical Inference of Weibull Distribution*, Chinese Journal of Applied Probability and Statistics, 8(1992)(4), 357–364.

[11] H. Xu, Y. F. Zhu and C. L. Chen, *Research on Battlefield Damaged Armored Materiel Maintenance Man Hour Simulation and Its Distribution*, Journal of System Simulation, 18(2006)(10), 2945–2951.

[12] K. Yamaoka, T. Nakagawa and T. Uno, *Statistical moments in pharmacokinetics,* J. Pharmacokin Biopharm, 6(1978)6, 547–558.

[13] H. W. Zhou, *Mathematical Medicine,* Shanghai Science and Technology Press, 1983.