Thailand Statistician
July 2026; 24(3): 530-550
http://statassoc.or.th
Contributed paper

# The New Generalized 2Kth-Order Equilibrium Model: Estimation and Characterization

**Nuzhat Ahad [a], S.P. Ahmad [a] and J.A. Reshi*[b]**
[a] Department of Statistics,University of Kashmir, Srinagar, J&K , India
[b] Department of Statistics,Govt. Degree College, Pulwama, J&K , India

*Corresponding author; e-mail: drjareshi@gmail.com

Received: 4 November 2024
Revised: 4 March 2026
Accepted: 9 April 2026

## Abstract

This manuscript introduces a generalization of familiar probability distributions such as Maxwell, Rayleigh, Half-normal, Chi-square, Gamma, Weibull and some related distributions of Rayleigh and Maxwell-Boltzmann distribution. It utilizes the concept of a weighted probability distribution, applying higher-order weighting to more precisely capture and represent complex real-life scenarios. The introduced model is termed as the Generalized equilibrium Maxwell-Boltzmann distribution (GEMBD). Various structural properties and characterizations of the newly proposed model have been derived. The ordering properties of the proposed model are analyzed and compared with those of the base distribution. Parameters are estimated via the maximum likelihood estimation (MLE) method. A simulation study is conducted using the AndersonDarling test statistic to assess the asymptotic normality of the MLEs. Additionally, the behaviors of bias and mean square error are observed with the increase in sample size. The applications of new distribution are illustrated through its fitting to two different real life datasets. Ultimately, a comparison is conducted among GEMBD and its sub-models regarding their fit using Information Criterion tools.

**Keywords:** Generalized equilibrium, Maxwell-Boltzmann distribution, Maximum likelihood estimation, AndersonDarling test statistic, Information criterion tools.

## 1. Introduction

The difficulty of accurately modeling real-world phenomena with standard probability distributions has led to the development of numerous innovative distributions. In many cases, existing distributions are insufficient for accurately representing specific real-life occurrences, necessitating further advancements and modifications, prompting researchers to continually refine and enhance these models. This research has driven the expansion and generalization of probability distributions, creating more adaptable and comprehensive models capable of capturing the intricate nuances of real-world data. A key advancement in this area has been the improvement of tail behavior representation by incorporating additional parameters into well-known distributions, thus enhancing their relevance to real-world datasets. For example, Maxwell's velocity model introduced a mathematical form of the Maxwell distribution to describe the distribution of particle velocities in thermal equilibrium. The Maxwell distribution (Maxwell (1868)), characterized by its unimodal and leptokurtic curves, accounts for the fact that molecules move at varying speeds and explains key properties of gases.

Recent studies have proposed several generalizations of the Maxwell-Boltzmann distribution, which have been rigorously validated. One such extension, the Generalized Maxwell-Boltzmann distribution (GMBD) is introduced by Voda (2009), and its probability density function (PDF) is defined as follows:

$$f(y) = \frac{\eta}{2^{\frac{1}{2\eta}}\Gamma(1+\frac{1}{2\eta})}\zeta^{-(2+\frac{1}{\eta})}y^{2\eta}e^{\frac{-y^{2\eta}}{2\zeta^2}}\ ; \quad y>0, \zeta>0, \eta>0 \tag{1}$$

Weighted distribution theory provides a robust framework for addressing challenges in model formulation and data interpretation. Rao (1965) emphasized the importance of statistical methodology in extracting relevant information from a sample to answer specific questions about the parent population. Misidentifying the population can lead to invalid inferences. To mitigate this issue, Rao (1965) introduced the concept of weighted distributions, which serve as a method of adjustment applicable across various contexts. These weighted distributions have demonstrated significant effectiveness in statistical modeling and prediction, particularly when standard distributions fall short. Recent contributions to this field include the works of Yu et al. (2020), Reshi et al. (2021), Bhat and Ahmad (2022), Shanker et al. (2023) and Ahad et al. (2024). The theory has been further generalized to higher orders, resulting in a hierarchy of distributions. A significant advancement was made with the introduction of higher-order equilibrium distributions by Harkness and Shantaram (1969). Recent applications have explored innovative concepts, including new partial orderings and aging properties, as demonstrated in the studies by Anas et al. (2023), introduces the Kth-order weighted equilibrium Weibull distribution, examining and comparing the aging and ordering properties between the baseline distribution and its equilibrium counterpart.This contribution offers valuable insights into the evolution of these distributions over time and underscores their practical applications in various fields.

In recent years, equilibrium distributions and their properties have attracted significant attention, particularly in reliability modeling and analysis. Understanding the relationship between these distributions and various reliability concepts is crucial for accurately describing lifetime models. Ongoing research continues to deepen our understanding of equilibrium distributions, refining methodologies for their effective utilization in real-world applications. In this manuscript, the GMBD of Voda (2009) is generalized by applying the higher-order weighted generalized method called Generalized even power weighted probability technique , previously used in Ahad et al. (2026). The motivation for adopting this technique is to introduce greater flexibility in the distributional shape, allowing improved modeling of various hazard rate shapes, thereby enhancing the models ability to fit complex lifetime data. The corresponding probability density function (PDF) is given as follows.

$$f_{w^{2k}}(y) = \frac{(w(y))^{2k}f(y)}{E[(w(y))^{2k}]};\ -\infty<y<\infty, k\in R, \tag{2}$$

where, $w(y)$ is weight function and $E[(w(y))^{2k}]<\infty$.

## 2. Derivation of GEMBD

The PDF of GEMBD is obtained by taking the weight function $w(y)=y$ and the PDF (1) in the basic definition of Generalized even power weighted probability technique (2); we have the following PDF of GEMBD is given by:

$$f_k(y) = \frac{\eta y^{2(\eta+k)}e^{-\frac{y^{2\eta}}{2\zeta^2}}}{2^{\frac{2k+1}{2\eta}}\zeta^{\frac{2(k+\eta)+1}{\eta}}\Gamma(\frac{2(k+\eta)+1}{2\eta})}\ ; \quad y>0, \zeta>0, \eta>0, k\in R, (k+\eta)>-0.5. \tag{3}$$

The corresponding cumulative distribution function of GEMBD is given by:

$$F_k(y) = 1-\frac{\Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{y^{2\eta}}{2\zeta^2})}{\Gamma(\frac{2(k+\eta)+1}{2\eta})}\ ; \quad y>0, \zeta>0, \eta>0, k\in R, (k+\eta)>-0.5. \tag{4}$$

The first order derivative of the density $f_k(y)$ is given by:

$$f_k'(y) = \frac{\eta y^{2(k+\eta)-1} e^{-\frac{y^{2\eta}}{2\zeta^2}}}{2^{\frac{2k+1}{2\eta}} \zeta^{\frac{2(k+\eta)+1}{\eta}} \Gamma(\frac{2(k+\eta)+1}{2\eta})} \left(2(k+\eta) - \frac{\eta y^{2\eta}}{\zeta^2}\right) ; \quad y>0, \zeta>0, \eta>0, k \in R, (k+\eta) > -0.5. \tag{5}$$

- For $(k+\eta) \leq 0$ , it is clear that $f_k'(y)$ is decreasing for all $y > 0$. Therefore the mode of GEMBD is zero for $(k+\eta) \leq 0$.
- For $(k+\eta) > 0$, the $f_k'(y)$ is increasing for $y > \zeta^{\frac{1}{\eta}}(\frac{2(k+\eta)}{\eta})^{\frac{1}{2\eta}}$ and is decreasing for $y < \zeta^{\frac{1}{\eta}}(\frac{2(k+\eta)}{\eta})^{\frac{1}{2\eta}}$ and maximum at $y = \zeta^{\frac{1}{\eta}}(\frac{2(k+\eta)}{\eta})^{\frac{1}{2\eta}}$. Therefore mode of GEMBD is $y = \zeta^{\frac{1}{\eta}}(\frac{2(k+\eta)}{\eta})^{\frac{1}{2\eta}}$.

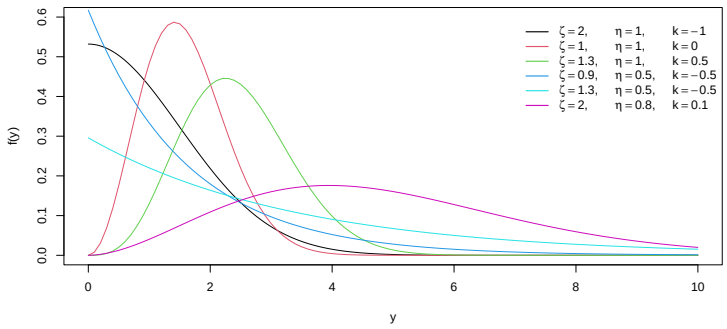



**Figure 1** The PDF curves of GEMBD for different parameter values

| Models | Abbreviations | Citations |
|---|---|---|
| Half normal distribution | HND | Leone et al. (1961),Johnson et al. (1995) |
| Rayleigh distribution | RD | Rayleigh (1880) |
| Maxwell-Boltzmann distribution | MBD | Maxwell (1868) |
| Length baised Maxwell distribution | LBMD | Modi and Gill (2015),Saghir et al. (2017) |
| Length baised weighted rayleigh distribution | LBWRD | Al-Khadim and Hussein (2014) |
| Area baised Maxwell distribution | ABMD | Sharma et al. (2018) |
| Chi-Square distribution | chi | Helmert (1876),Pearson (1900) |
| Weibull distribution | WD | Weibull (1939) |
| Exponential distribution | ED | |
| Standard gamma distribution | SGD | |

**Table 1** The sub-models of Generalized equilibrium Maxwell-Boltzmann distribution

| Sub-Models | Parameter restriction | PDF |
|---|---|---|
| HND | $k=-1,\eta=1$ | $f(y)=\frac{\sqrt{2}e^{-\frac{y^2}{2\zeta^2}}}{\sqrt{\pi}\zeta}$ |
| RD | $k=-\frac{1}{2},\eta=1$ | $f(y)=\frac{ye^{-\frac{y^2}{2\zeta^2}}}{\zeta^2}$ |
| MBD | $k=0,\eta=1$ | $f(y)=\sqrt{\frac{2}{\pi}}\zeta^{-3}y^2e^{\frac{-y^2}{2\zeta^2}}$ |
| LBMD | $k=\frac{1}{2},\eta=1$ | $f(y)=\frac{y^3e^{\frac{-y^2}{2\zeta^2}}}{2\zeta^4}$ |
| LBWRD | $k=\frac{1}{2},\eta=1,\zeta=\frac{1}{\zeta'}$ | $f(y)=\frac{2y^3\eta^2e^{\frac{-\eta y^2}{\zeta'}}}{(\zeta')^2}$ |
| ABMD | $k=1,\eta=1$ | $f(y)=\frac{y^4\zeta^{-\frac{5}{2}}e^{\frac{-\zeta y^2}{2}}}{2^{\frac{3}{2}}\Gamma(\frac{5}{2})}$ |
| ED | $k=-\frac{1}{2},\eta=\frac{1}{2},\zeta^2=\frac{\zeta'}{2}$ | $f(y)=\frac{1}{\zeta'}e^{-\frac{y}{\zeta'}}$ |
| Chi | $k=\frac{n}{4}-1,\eta=\frac{1}{2},\zeta=1$ | $f(y)=\frac{y^{\frac{n}{2}-1}e^{-\frac{y}{2}}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}$ |
| SGD | $k=\frac{k}{2}-1,\eta=\frac{1}{2},\zeta^2=\frac{1}{2}$ | $f(y)=\frac{y^{k-1}e^{-y}}{\Gamma(k)}$ |
| WD | $k=-\frac{1}{2},\eta=\frac{\eta'}{2},2\zeta^2=(\zeta')^{\eta'}$ | $f(y)=\frac{\eta' y^{\eta'-1}e^{-(\frac{y}{\zeta'})^{\eta'}}}{(\zeta')^{\eta'}}$ |

## 3. Reliability Analysis

### 3.1. Reliability function and hazard rate function of GEMBD

The reliability function $(R_k(Y))$ and hazard rate function $(h_k(Y))$ of GEMBD is given respectively by

$$R_k(y)=\frac{\Gamma(\frac{2(\eta+k)+1}{2\eta},\frac{y^{2\eta}}{2\zeta^2})}{\Gamma(\frac{2(k+\eta)+1}{2\eta})};\quad y>0,\zeta>0,\eta>0,k\in R,(k+\eta)>-0.5. \tag{6}$$

$$h_k(y)=\frac{\eta y^{2(k+\eta)}e^{\frac{-y^{2\eta}}{2\zeta^2}}}{2^{\frac{2k+1}{2\eta}}\zeta^{\frac{2(k+\eta)+1}{\eta}}\Gamma(\frac{2(k+\eta)+1}{2\eta},\frac{y^{2\eta}}{2\zeta^2})};\quad y>0,\zeta>0,\eta>0,k\in R,(k+\eta)>-0.5. \tag{7}$$

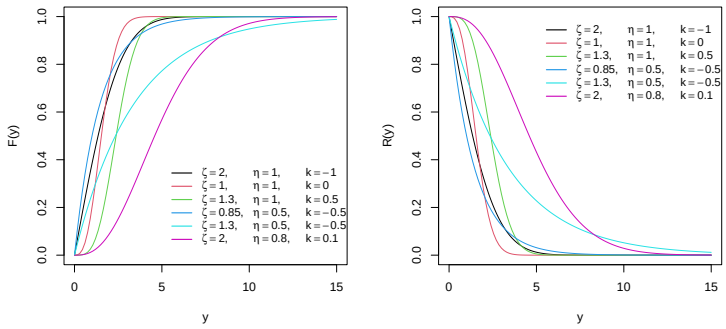



**Figure 2** The cumulative distribution function and reliability function of GEMBD for selected parameter values

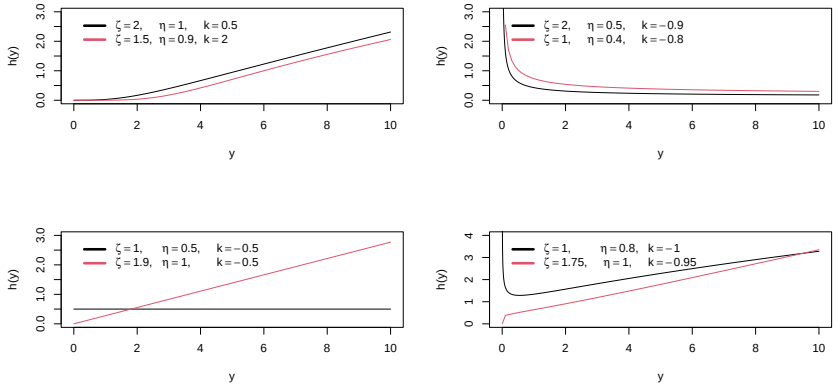



**Figure 3** The hazard rate function curves of GEMBD for selected value of parameters

Figure 3 describes the behaviour of hazard rate function of GEMBD for different choice of parameters and it depicts that the hazard rate function of GEMBD is increasing, decreasing, constant and decreasing- increasing.

- For $k < -0.5$ , $h_k(y)$ is decreasing and decreasing-increasing and while as, for $k > 0.5$, $h_k(y)$ is increasing.
- For $-0.5 \leq k < 0.5$, $h_k(y)$ is constant and increases linearly.

### 3.2. Ordering properties

The ratio of the densities of base model random variable (i.e. GMBD $Y_{GMBD}$) and its corresponding equilibrium version ( i.e. GEMBD $Y_{GEMBD}$ ) is given by

$$r(y) = \frac{2^{\frac{k}{\eta}} \zeta^{\frac{2k}{\eta}} y^{-2k} \Gamma(\frac{2(k+\eta)+1}{2\eta})}{\Gamma \frac{2\eta+1}{2\eta}} \tag{8}$$

and the first order derivative of $r(y)$ is given by

$$r'(y) = \frac{2^{\frac{k}{\eta}}\zeta^{\frac{2k}{\eta}}\Gamma(\frac{2(k+\eta)+1}{2\eta})}{\Gamma\frac{2\eta+1}{2\eta}}(-2k)y^{-2k-1} \tag{9}$$

The expression indicates that $r'(y)$ is positive for $k < 0$, negative for $k > 0$, and zero for $k = 0$. Consequently, it follows that,

- For $k < 0$, $Y_{GEMBD}$ exhibits a lower likelihood ratio compared to $Y_W$ ($Y_{GEMBD} \leq_{lr} Y_{GMBD}$). Based on the chain of implications within various stochastic orders, it can be confidently concluded that $Y_{GEMBD} \leq_h Y_{GMBD}$ and $Y_{GEMBD} \leq_{mrl} Y_{GMB}$. for $k > 0$, converse is true.
  Where $h \longrightarrow$ hazard rate and $mrl \longrightarrow$ mean residual life.

The validity of these properties between $Y_{GEMBD}$ and $Y_{GMBD}$ has been confirmed at predefined values of $\zeta$ and $k$, as detailed in Table 2. This analysis encompasses three cases: $k < 0$, $k > 0$, and $k = 0$. And The data presented in Table 2 and Figure 4 distinctly support the aforementioned ordering relations.

**Table 2** Ordering relations between GEMBD and Generalized Maxwell-Boltzmann distribution

| Cases | k | $Y = 10, \zeta = 3, \eta = 1$ | | | $Y = 10, \zeta = 5, \eta = 1.5$ | | |
|---|---|---|---|---|---|---|---|
| | | $r'(y)$ | h(y) | | $r'(y)$ | h(y) | |
| | | | $Y_{GMBD}$ | $Y_{GEMBD}$ | | $Y_{GMBD}$ | $Y_{GEMBD}$ |
| $k < 0$ | -1 | 0.7407 | 1.0255 | 1.1982 | 0.7425 | 5.9046 | 6.0955 |
| | -0.5 | 0.0887 | 1.0255 | 1.1111 | 0.1010 | 5.9046 | 6.0000 |
| | -0.1 | 0.0079 | 1.0255 | 1.0425 | 0.0082 | 5.9046 | 5.9237 |
| $k = 0$ | 0 | 0 | 1.0255 | 1.0255 | 0 | 5.9046 | 5.9046 |
| $k > 0$ | 0.1 | -0.0057 | 1.0255 | 1.0086 | -0.0054 | 5.9046 | 5.8856 |
| | 0.5 | -0.0160 | 1.0255 | 0.9416 | -0.0124 | 5.9046 | 5.8094 |
| | 1 | -0.0180 | 1.0255 | 0.8596 | -0.0101 | 5.9046 | 5.7143 |

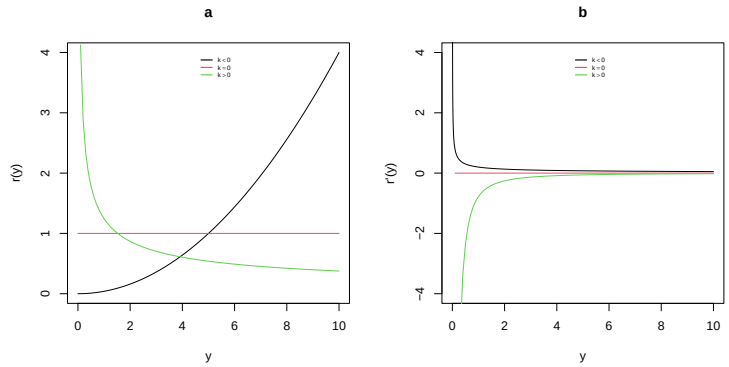



**Figure 4** (a) Ratio of density function of GMBD and GEMBD, (b) First order derivative of ratio of density function of GMBD and GEMBD

### 3.3. Reverse hazard rate function

The reverse hazard rate function $r_k(y)$ of GEMBD is given by

$$r_k(y) = \frac{\eta y^{2(k+\eta)} e^{-\frac{y^{2\eta}}{2\zeta^2}} \zeta^{-\frac{(2(k+\eta)+1)}{\eta}}}{2^{\frac{2k+1}{2\eta}} \left[\Gamma(\frac{2(k+\eta)+1}{2\eta}) - \Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{y^{2\eta}}{2\zeta^2})\right]}; \quad y > 0, \zeta > 0, \eta > 0, k \in R, (k+\eta) > -0.5. \tag{10}$$

## 4. Structural Properties of GEMBD

In this section, some statistical properties of GEMBD are discussed.

### 4.1. Moments

In this subsection we compute the $r^{th}$ moment, mean, variance, coefficient of variation (CV), coefficient of skewness (CS) and coefficient of kurtosis (CK) of GEMBD and is given respectively by

$$E(Y^r) = (\sqrt{2}\zeta)^{\frac{r}{\eta}} \frac{\Gamma(\frac{2(k+\eta)+r+1}{2\eta})}{\Gamma(\frac{2(k+\eta)+1}{2\eta})} \tag{11}$$

$$E(Y) = (\sqrt{2}\zeta)^{\frac{1}{\eta}} \frac{\Gamma(\frac{2(k+\eta)+2}{2\eta})}{\Gamma(\frac{2(k+\eta)+1}{2\eta})} \tag{12}$$

$$V(Y) = \frac{(2\zeta^2)^{\frac{1}{\eta}}}{\left(\Gamma(\frac{2(k+\eta)+1}{2\eta})\right)^2} \left[\Gamma(\frac{2(k+\eta)+3}{2\eta})\Gamma(\frac{2(k+\eta)+1}{2\eta}) - \left(\Gamma(\frac{2(k+\eta)+2}{2\eta})\right)^2\right] \tag{13}$$

$$CV = \frac{\left[\Gamma(\frac{2(k+\eta)+1}{2\eta})\Gamma(\frac{2(k+\eta)+3}{2\eta}) - \left(\Gamma(\frac{2(k+\eta)+2}{2\eta})\right)^2\right]^{\frac{1}{2}}}{\Gamma(\frac{2(k+\eta)+2}{2\eta})} \tag{14}$$

$$CS = \frac{\Gamma(\frac{2(k+\eta)+4}{2\eta})\left(\Gamma(\frac{2(k+\eta)+1}{2\eta})\right)^2 - 3\Gamma(\frac{2(k+\eta)+1}{2\eta})\Gamma(\frac{2(k+\eta)+2}{2\eta})\Gamma(\frac{2(k+\eta)+3}{2\eta}) + 2\left(\Gamma(\frac{2(k+\eta)+2}{2\eta})\right)^3}{\left[\Gamma(\frac{2(k+\eta)+1}{2\eta})\Gamma(\frac{2(k+\eta)+3}{2\eta}) - \left(\Gamma(\frac{2(k+\eta)+2}{2\eta})\right)^2\right]^{\frac{3}{2}}} \tag{15}$$

$$CK = \frac{H}{\left[\Gamma(\frac{2(k+\eta)+1}{2\eta})\Gamma(\frac{2(k+\eta)+3}{2\eta}) - \left(\Gamma(\frac{2(k+\eta)+3}{2\eta})\right)^2\right]^2} \tag{16}$$

where

$$\begin{aligned} H =& \Gamma(\frac{2(k+\eta)+5}{2\eta})\left(\Gamma(\frac{2(k+\eta)+1}{2\eta})\right)^3 - 4\Gamma(\frac{2(k+\eta)+2}{2\eta})\Gamma(\frac{2(k+\eta)+4}{2\eta})\left(\Gamma(\frac{2(k+\eta)+1}{2\eta})\right)^2 \\ &+ 6\left(\Gamma(\frac{2(k+\eta)+2}{2\eta})\right)^2 \Gamma(\frac{2(k+\eta)+3}{2\eta})\Gamma(\frac{2(k+\eta)+1}{2\eta}) - 3\left(\Gamma(\frac{2(k+\eta)+2}{2\eta})\right)^4 \end{aligned}$$

### 4.2. Incomplete moments and conditional moments

The incomplete moments display the graphical structure of a distributions moments, which is helpful in various fields such as econometrics, finance, and reliability. The $n^{th}$ incomplete moments

$(I_k(n,t))$ of GEMBD is given by

$$I_k(n,t) = \int_0^t y^n f_k(y)dy$$

$$I_k(n,t) = \frac{(\sqrt{2}\zeta)^{\frac{n}{2\eta}}\gamma(\frac{2(k+\eta)+2}{2\eta}, \frac{t^{2\eta}}{2\zeta^2})}{\Gamma(\frac{2(k+\eta)+1}{2\eta})} \tag{17}$$

The $n^{th}$ conditional moment of GEMBD is given by

$$E(T^n|T>t) = \frac{\int_t^\infty y^n f_k(y)dy}{R(t)}$$

where, $\int_t^\infty y^n f_k(y)dy = \frac{(\sqrt{2}\zeta)^{\frac{n}{\eta}}\Gamma(\frac{n+2(k+\eta)+1}{2\eta}, \frac{t^{2\eta}}{2\zeta^2})}{\Gamma(\frac{2(k+\eta)+1}{2\eta})}$.

$$E(T^n|T>t) = \frac{(\sqrt{2}\zeta)^{\frac{n}{\eta}}\Gamma(\frac{n+2(k+\eta)+1}{2\eta}, \frac{t^{2\eta}}{2\zeta^2})}{\Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{t^{2\eta}}{2\zeta^2})} \tag{18}$$

### 4.3. Mean residual life and mean waiting time

Mean residual life $\mu_k(t)$ of GEMBD is given by:

$$\mu_k(t) = \frac{1}{R_k(t)}\left[E(t) - \int_0^t y f_k(y)dy\right] - t$$

where, $\int_0^t y f_k(y)dy = \frac{(\sqrt{2}\zeta)^{\frac{1}{\eta}}\gamma(\frac{2(k+\eta)+2}{2\eta}, \frac{t^{2\eta}}{2\zeta^2})}{\Gamma(\frac{2(k+\eta)+1}{2\eta})}$

$$\mu_k(t) = \frac{(\sqrt{2}\zeta)^{\frac{1}{\eta}}}{\Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{t^{2\eta}}{2\zeta^2})}\left[\Gamma(\frac{2(k+\eta)+2}{2\eta}) - \gamma(\frac{2(k+\eta)+2}{2}, \frac{t^{2\eta}}{2\zeta^2})\right] - t \tag{19}$$

The mean waiting time $\hat{\mu}_k(t)$ of GEMBD is given by:

$$\hat{\mu}_k(t) = t - \frac{1}{F_k(t)}\int_0^t y f_k(y)dy$$

$$\hat{\mu}_k(t) = t - \frac{(\sqrt{2}\zeta)^{\frac{1}{\eta}}}{\Gamma(\frac{2(k+\eta)+1}{2\eta}) - \Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{t^{2\eta}}{2\zeta^2})}\gamma(\frac{2(k+\eta)+2}{2\eta}, \frac{t^{2\eta}}{2\zeta^2}) \tag{20}$$

### 4.4. Entropy measures of GEMBD

Entropy measures provide a quantification of uncertainty or randomness within the system. We have calculated the expressions for Renyi entropy, and is given as:

$$H_R(\delta) = \frac{1}{1-\delta}log\left[\frac{\eta^{\delta-1}(\sqrt{2})^{(\delta-1)(\frac{2\eta-1}{\eta})}\zeta^{\frac{1}{\eta}(1-\delta)}}{\delta^{\frac{2\delta(k+\eta)+1}{2\eta}}\left(\Gamma(\frac{2(k+\eta)+1}{2\eta}\right)^\delta}\Gamma(\frac{2\delta(k+\eta)+1}{2\eta})\right] \tag{21}$$

### 4.5. Odds ratio and Mills ratio

The odds ratio $O_k(y)$ and Mills ratio $M_k(y)$for GEMBD distribution is given respectively as

$$O_k(y) = \frac{F_k(y)}{f_k(y)} = \frac{2^{\frac{2k+1}{2\eta}}\zeta^{\frac{2(k+\eta)+1}{\eta}}\left[\Gamma(\frac{2(k+\eta)+1}{2\eta}) - \Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{y^{2\eta}}{2\zeta^2})\right]}{\eta y^{2(k+\eta)}e^{-\frac{y^{2\eta}}{2\zeta^2}}} \tag{22}$$

$$M_k(y) = \frac{R_k(y)}{f_k(y)} = \frac{2^{\frac{2k+1}{2\eta}}\zeta^{\frac{2(k+\eta)+1}{\eta}}\Gamma(\frac{2(k+\eta)+1}{2}, \frac{y^{2\eta}}{2\zeta^2})}{\eta y^{2(k+\eta)}e^{-\frac{y^{2\eta}}{2\zeta^2}}} \tag{23}$$

### 4.6. Lorenz inequality and Bonferoni inequality

The Bonferroni and Lorenz curve for a random variable Y following GEMBD are respectively given by

$$L(t) = \frac{I_k(t)}{E(t)} = \frac{\gamma(\frac{2(k+\eta)+2}{2}, \frac{t^{2\eta}}{2\zeta^2})}{\Gamma(\frac{2(k+\eta)+2}{2\eta})} \tag{24}$$

$$B(t) = \frac{L(t)}{F(t)} = \frac{\Gamma(\frac{2(k+\eta)+1}{2\eta})\gamma(\frac{2(k+\eta)+2}{2\eta}, \frac{t^{2\eta}}{2\zeta^2})}{\Gamma(\frac{2(k+\eta)+2}{2\eta})\left[\Gamma(\frac{2(k+\eta)+1}{2\eta} - \Gamma(\frac{2(k+\eta)+2}{2\eta}, \frac{t^{2\eta}}{2\zeta^2})\right]} \tag{25}$$

### 4.7. Order statistics

Let us assume that the random sample $Y_1, Y_2, Y_3..., Y_n$ from the GEMBD, with PDF $(f_k(y))$ and CDF $(F_k(y))$ Let $Y_{(1)}, Y_{(2)}, Y_{(3)}, ..., Y_{(n)}$ denote the corresponding order statistics, then the PDF and cumulative distribution function (CDF) of the $r^{th}(r = 1, 2, ...n)$ order statistic are respectively, given by

$$f_r(y) = \frac{n!}{(r-1)!(n-r)!}[F_k(y)]^{r-1} f_k(y)[(1 - F_k(y))]^{n-r}. \tag{26}$$

and

$$F_r(y) = \sum_{j=r}^{n} {}^nC_j F_k^J(y)\,[1 - F_k(y)]^{n-j} \tag{27}$$

$$\begin{aligned} f_r(y) =& \frac{\eta n! y^{2(k+\eta)} \zeta^{-\frac{(2(k+\eta)+1)}{\eta}} e^{-\frac{y^{2\eta}}{2\zeta^2}}}{(r-1)!(n-r)! 2^{\frac{2k+1}{2\eta}} (\Gamma(\frac{2(k+\eta)+1}{2\eta}))^n} \\ & * \left[\Gamma(\frac{2(k+\eta)+1}{2\eta}) - \Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{y^{2\eta}}{2\zeta^2})\right]^{r-1} \left[\Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{y^{2\eta}}{2\zeta^2})\right]^{n-r} \end{aligned} \tag{28}$$

and

$$F_r(y) = \sum_{j=r}^{n} \binom{n}{j} \left[\frac{\Gamma(\frac{2(k+\eta)+1}{2\eta}) - \Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{y^{2\eta}}{2\zeta^2})}{\Gamma(\frac{2(k+\eta)+1}{2})}\right]^j \left[\frac{\Gamma(\frac{2(k+\eta)+1}{2}, \frac{y^{2\eta}}{2\zeta^2})}{\Gamma(\frac{2(k+\eta)+1}{2\eta})}\right]^{n-j} \tag{29}$$

## 5. Characterizations of GEMBD

This section deals with various characterizations of GEMBD. These characterizations are based on: Coefficient of variation; a simple relationship between two truncated moments; the hazard function; and reverse hazard function.

### 5.1. Characterization based on coefficient of variation

**Proposition 1** *Let $Y_1, Y_2, Y_3, ..., Y_n$ be a random sample of size $n$ drawn from GEMBD, then the square of sample coefficient of variation is asymptotically unbiased estimator of the square of population coefficient of variation.*

*Mathematically,* $\lim_{n\to\infty} E[\frac{{S_n}^2}{\bar{Y_n}^2}] = lim_{n\to\infty} E[\frac{S_n}{\bar{Y_n}}]^2 = [\frac{\sigma}{\mu}]^2$. *Where* $\bar{Y_n} = \mu$ *and* ${S_n}^2 = \frac{\sigma^2}{n}$ *is mean and variance respectively.*

**Proof:**

$$E(\bar{y_n}^2) = \frac{(\sqrt{(}2)\zeta)^{\frac{2}{\eta}}}{n\left(\Gamma(\frac{2(k+\eta)+2)}{2\eta}\right)^2}\left[\Gamma(\frac{2(k+\eta)+3)}{2\eta}\Gamma(\frac{2(k+\eta)+1)}{2\eta} + (n-1)\left(\Gamma(\frac{2(k+\eta)+2)}{2\eta}\right)^2\right] \tag{30}$$

$$E({S_n}^2) = \sigma^2 = \frac{(\sqrt{(}2)\zeta)^{\frac{2}{\eta}}}{\left(\Gamma(\frac{2(k+\eta)+2)}{2\eta}\right)^2}\left[\Gamma(\frac{2(k+\eta)+3)}{2\eta}\Gamma(\frac{2(k+\eta)+1)}{2\eta} - \left(\Gamma(\frac{2(k+\eta)+2)}{2\eta}\right)^2\right]. \tag{31}$$

Now we can write,

$$E({S_n}^2) = E\left[\frac{{S_n}^2}{\bar{Y_n}^2}\bar{Y_n}^2\right]$$
$$E({S_n}^2) = E\left[\frac{{S_n}^2}{\bar{Y_n}^2}\right]E\left[\bar{Y_n}^2\right]$$
$$\frac{E\left[{S_n}^2\right]}{E\left[\bar{Y_n}^2\right]} = E\left[\frac{S_n}{\bar{Y_n}}\right]^2.$$

Using (30) and (31), we obtain

$$E\left[\frac{S_n}{\bar{Y_n}}\right]^2 = \frac{\left[\Gamma(\frac{2(k+\eta)+3)}{2\eta}\Gamma(\frac{2(k+\eta)+1)}{2\eta} - \left(\Gamma(\frac{2(k+\eta)+2)}{2\eta}\right)^2\right]}{\left[\frac{1}{n}\Gamma(\frac{2(k+\eta)+3)}{2\eta}\Gamma(\frac{2(k+\eta)+1)}{2\eta} + (1-\frac{1}{n})\left(\Gamma(\frac{2(k+\eta)+2)}{2\eta}\right)^2\right]}$$

taking $\lim_n \to \infty$ on both sides we get

$$\lim_{n\to\infty} E\left[\frac{S_n}{\bar{Y_n}}\right]^2 = \frac{\left(\sqrt{(}\Gamma(\frac{2(k+\eta)+3}{2\eta})\Gamma(\frac{2(k+\eta)+1}{2\eta}) - (\Gamma(\frac{2(k+\eta)+2}{2\eta})^2)\right)^2}{\left(\Gamma(\frac{2(k+\eta)+2)}{2\eta}\right)^2}$$

$$\lim_{n\to\infty} E\left[\frac{S_n}{\bar{Y_n}}\right]^2 = \left(\frac{\sigma}{\mu}\right)^2. \tag{32}$$

### 5.2. Characterization based on truncated moments

**Proposition 2** *Let $Y : \Omega \to (0,\infty)$ be a continuous random variable. Let $b_1(y) = y^{-(2k+1)}$ and $b_2(y) = b_1(y)e^{-\frac{y^{2\eta}}{2\zeta^2}}$ for $y > 0$ then the random variable $Y$ has PDF (3) iff the function $A(y) = \frac{e^{-\frac{y^{2\eta}}{2\zeta^2}}}{2}$.*

**Proof:** Suppose the random variable $Y$ has PDF (3).

$$E[b_1(y)|Y \geq y] = q\zeta^2 e^{-\frac{y^{2\eta}}{2\zeta^2}}$$
$$E[b_2(y)|Y \geq y] = \frac{q}{2}\zeta^2 e^{-\frac{2y^{2\eta}}{2\zeta^2}} = E[b_1(y)|Y \geq y]A(y)$$

where $q$ is appropriate constant. Conversely, if $A(y)$ is of the above form, then thus from Theorem 1 $F$ is uniformly distributed by $b_1(y), b_1(y)$ such that

$$F(y) = \int_0^\infty C|\frac{A^{'}(y)}{A(y)b_1(y) - b_2(y)}|e^{-S(y)}dy \tag{33}$$
$$S^{'}(y) = \frac{A^{'}(y)b_(y)}{A(y)b_1(y) - b_2(y)} = \frac{\eta y^{2\eta-1}}{\zeta^2}$$
$$S(y) = \frac{y^{2\eta}}{2\zeta^2}$$

Using $S(y)$ and normalizing constant C value in (33) we get the CDF of GEMBD.

### 5.3. Characterization based on hazard rate function (HRF)

**Proposition 3** *Let $Y : \Omega \to (0, \infty)$ be a continuous random variable. The PDF of $Y$ is (3) iff its hazard rate function $h(y)$ satisfies the differential equation,*

$$h^{'}(y) - \frac{2(k+\eta)}{y}h(y) = \frac{-\eta}{q\zeta^2}y^{2(k+\eta)-2\eta-1}e^{-\frac{y^{2\eta}}{2\zeta^2}} \tag{34}$$

**Proof:** If $Y$ has PDF (3) we get

$$h^{'}(y) - \frac{2(k+\eta)}{y}h(y) = \frac{-\eta}{q\zeta^2}y^{2(k+\eta)-2\eta-1}e^{-\frac{y^{2\eta}}{2\zeta^2}}.$$

Conversely, if HRF satisfies the differential equation above then we have

$$\frac{d}{dy}y^{-2(k+\eta)}h(y) = \frac{1}{q}\frac{d}{dy}e^{-\frac{y^{2\eta}}{2\zeta^2}}$$
$$h(y) = \frac{1}{q}y^{2(k+\eta)}e^{-\frac{y^{2\eta}}{2\zeta^2}}$$

which is HRF of GEMBD distribution.

### 5.4. Characterization based on reverse hazard rate function (RHRF)

**Proposition 4** *Let $Y : \Omega \to (0, \infty)$ be a continuous random variable. The PDF of $Y$ is (3) iff its reverse hazard rate $r(y)$ satisfies the differential equation*

$$r^{'}(y) - \frac{2(k+\eta)}{y}r(y) = \frac{\eta}{q}y^{2(k+\eta)}\frac{d}{dy}\frac{e^{-\frac{y^{2\eta}}{2\zeta^2}}}{\Gamma(\frac{2(k+\eta)+1}{2\eta}) - \Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{y^{2\eta}}{2\zeta^2})}. \tag{35}$$

**Proof:** If $Y$ has PDF (3) we get

$$r^{'}(y) - \frac{2(k+\eta)}{y}r(y) = \frac{\eta}{q}y^{2(k+\eta)}\frac{d}{dy}\frac{e^{-\frac{y^{2\eta}}{2\zeta^2}}}{\Gamma(\frac{2(k+\eta)+1}{2\eta}) - \Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{y^{2\eta}}{2\zeta^2})}$$

Conversely, if RHRF satisfies the differential equation above then we have

$$\frac{d}{dy}y^{-2(k+\eta)}r(y) = \frac{\eta}{q}\frac{d}{dy}\frac{e^{n-\frac{y^{2\eta}}{2\zeta^2}}}{\Gamma(\frac{2(k+\eta)+1}{2\eta}) - \Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{y^{2\eta}}{2\zeta^2})}$$
$$r(y) = \frac{\eta}{q}y^{2(k+\eta)}\frac{e^{-\frac{y^{2\eta}}{2\zeta^2}}}{\Gamma(\frac{2(k+\eta)+1}{2\eta}) - \Gamma(\frac{2(k+\eta)+1}{2\eta}, \frac{y^{2\eta}}{2\zeta^2})}$$

which is RHRF of GEMBD.

### 5.5. Parametric estimation of GEMBD using maximum likelihood estimation technique

Let $Y_1, Y_2, Y_3, ..., Y_n$ be an observed sample taken from the $GEMBD(\zeta, \eta, k)$ with unknown parameters $\zeta, \eta$ and $k$, then the log-likelihood function can be written as

$$\log L(y,\zeta,k) = n\log\eta - n\left(\frac{2k+1}{2\eta}\right)\log 2 + 2(k+\eta)\sum_{i=1}^{n}\log(y_i) - \frac{\sum_{i=1}^{n} y_i^{2\eta}}{2\zeta^2} - \frac{n(2(k+\eta)+1)}{2\eta}\log\zeta - n\log\Gamma(\frac{2(k+\eta)+1}{2\eta}). \tag{36}$$

The values of $\hat{\zeta}, \hat{\eta}$ and $\hat{k}$ that maximize the log-likelihood function are called the maximum likelihood estimates of the parameters $\zeta, \eta$ and $k$.

The equations obtained on equating the first-order partial derivatives of $\log L(y, \zeta, k)$ with respect to $\zeta, \eta$ and $k$ to zero, are given as

$$\frac{\sum_{i=1}^{n}(y_i)^{2\eta}}{\zeta^3} - \frac{n(2(k+\eta)+1)}{\eta\zeta} = 0 \tag{37}$$

$$-\frac{n}{\eta}+2\sum_{i=1}^{n}\log(y_i)+\frac{n(2k+1)\log 2}{2\eta^2}-\frac{\sum_{i=1}^{n}(y_i)^{2\eta}\log(y_i)}{\zeta^2}+\frac{n(2k+1)}{\eta^2}\log(\zeta)+\frac{n(2k+1)}{2\eta^2}\psi(\frac{2(k+\eta)+1}{2\eta}) = 0 \tag{38}$$

$$-\frac{n\log 2}{\eta} + 2\sum_{i=1}^{n}\log(y_i) - 2\frac{n}{\eta}\log\zeta - \frac{n}{\eta}\psi(\frac{2(k+\eta)+1}{2\eta}) = 0 \tag{39}$$

$\psi(\cdot)$ is digamma function.

The associated Fisher information matrix is given by

$$I(\hat{\zeta},\hat{\eta},\hat{k};y) = -\begin{pmatrix} I_{\zeta\zeta} & I_{\zeta\eta} & I_{\zeta k} \\ I_{\eta\zeta} & I_{\eta\eta} & I_{\eta k} \\ I_{k\zeta} & I_{k\eta} & I_{kk} \end{pmatrix},$$

where

$$I_{\zeta\zeta} = \frac{\partial^2\log L}{\partial\zeta^2} = \frac{-3\sum_{i=1}^{n}(y_i)^{2\eta}}{\zeta^4} + \frac{n(2(k+\eta)+1)}{\eta\zeta^2}$$

$$I_{\eta\zeta} = \frac{\partial^2\log L}{\partial\eta\partial\zeta} = \frac{2\sum_{i=1}^{n}(y_i)^{2\eta}\log(y_i)}{\zeta^3} + \frac{n(k+1)}{\eta\zeta^2}$$

$$I_{\eta\eta} = \frac{\partial^2\log L}{\partial\eta^2} = -\frac{n}{\eta^2} - \frac{n(2k+1)\log 2}{2\eta^3} - \frac{2\sum_{i=1}^{n}(y_i)^{2\eta}(log(y_i))^2}{\zeta^2} - \frac{n(2k+1)}{\eta^3}\log\zeta - \frac{4nk}{\eta^3}\psi(\frac{2(k+\eta)+1}{2\eta}) - \frac{n}{2}(\frac{(2k+1)}{\eta^2})^2\psi'(\frac{2(k+\eta)+1}{2\eta})$$

$$I_{k\zeta} = \frac{\partial^2\log L}{\partial k\partial\zeta} = \frac{-2n}{\eta\zeta}$$

$$I_{k\eta} = \frac{\partial^2\log L}{\partial k\partial\eta} = \frac{n(\log 2)}{\eta^2} + \frac{2n}{\eta^2}\log(\zeta) + \frac{n(2k+1)}{2\eta^3}\psi'(\frac{2(k+\eta)+1}{2\eta}) + \frac{(n)}{\eta^2}\psi(\frac{2(k+\eta)+1}{2\eta})$$

$$I_{kk} = \frac{n}{\eta^2}\psi'(\frac{2(k+\eta)+1}{2\eta}).$$

Thus, the approximate variance-covariance matrix is given by

$$\begin{pmatrix} Var(\hat{\zeta}) & Cov(\hat{\zeta}\hat{\eta}) & Cov(\hat{\zeta}\hat{k}) \\ Cov(\hat{\eta}\hat{\zeta}) & Var(\hat{\eta}) & Cov(\hat{\eta}\hat{\zeta}) \\ Cov(\hat{k}\hat{\zeta}) & Cov(\hat{k}\hat{\eta}) & Var(\hat{k}) \end{pmatrix} = I^{-1}(\hat{\zeta}, \hat{\eta}, \hat{k}; y),$$

and the corresponding 95% confidence intervals of the parameters $(\zeta, \eta, k)$ are respectively, given by

$$\hat{\zeta} \pm Z_{\alpha/2}\sqrt{\text{Var}(\hat{\zeta})}$$

$$\hat{\eta} \pm Z_{\alpha/2}\sqrt{\text{Var}(\hat{\eta})}$$

$$\hat{k} \pm Z_{\alpha/2}\sqrt{\text{Var}(\hat{k})}$$

where, $Z_{\alpha/2}$ is the upper $\alpha/2$ point of the standard normal distribution and $\alpha$ is the significance level.

## 6. Simulation and Computation

In this section, a detailed simulation study is conducted to investigate the behavior of the maximum likelihood estimators (MLEs) for the parameters of GEMBD as the sample size increases. The methodology involves the following steps:

1. **Generation of random samples:** A total of 200 random samples of size $n$ are generated from GEMBD using the inverse cumulative distribution function method. In this method a sample is generated from a distribution by solving the equation $F(y, a) = P \sim U(0, 1)$ for $y$, at preassigned values of parametric space $a$ and at $n$ independent values of $P$. Since solving this equation for our proposed model analytically is computationally tedious, the uniroot function from the R package stats is employed to numerically find the root. By applying the uniroot function for each value of $\zeta, \eta$ and $\lambda$, we generated the desired sample from the GEMBD. The procedure are repeated 200 times to generate 200 random samples.

2. **Estimation of MLEs:** For each of these 200 samples, the MLEs for the parameters $(\hat{\zeta}_i, \hat{\eta}_i, \hat{k}_i)$ are computed, where $i = 1, 2, 3, \ldots, 200$.

3. **Calculating bias and MSE:** Using the generated estimates, we evaluate the bias and mean squared error (MSE) for each parameter $a = \zeta, \eta, k$ through the following formulas:

   - **Bias:**
   $$\text{Bias}_a = \frac{1}{200}\sum_{i=1}^{200}(\hat{a}_i - a),$$

   - **MSE:**
   $$\text{MSE}_a = \frac{1}{200}\sum_{i=1}^{200}(\hat{a}_i - a)^2,$$

4. **Anderson-Darling test for normality:** To test whether the MLEs converge to a normal distribution as the sample size increases, we apply the Anderson-Darling (AD) test. The AD-test statistic and the corresponding p-values are computed for various sample sizes. The results are visualized with trace plots that show how the test statistic and p-values change with increasing sample size, indicating the point at which the estimators approach normality.

5. **Determining asymptotic normality threshold:** A threshold sample size $n_0$ is identified, beyond which the MLEs begin to exhibit asymptotic normality. For sample sizes larger than $n_0$, new random samples are generated and repeated 100 times. The estimates for the vector $\hat{\mathbf{a}}_{\text{mle}} = [\hat{\zeta}_{\text{mle}}, \hat{\eta}_{\text{mle}}, \hat{k}_{\text{mle}}]^T$ are used to construct the empirical distribution $N_{\text{emp}}(\mathbf{a}_{\text{mle}}, I^{-1}(\mathbf{a}_{\text{mle}}))$. The mean vector and covariance matrix are given by:

$$\mathbf{a}_{\text{mle}} = \mathbb{E}[\hat{\zeta}_{\text{mle}}, \hat{\eta}_{\text{mle}}, \hat{k}_{\text{mle}}]^T,$$

$$I^{-1}(\mathbf{a}_{\text{mle}}) = \mathbb{E}[(\hat{\mathbf{a}}_{\text{mle}} - \mathbb{E}[\hat{\mathbf{a}}_{\text{mle}}])(\hat{\mathbf{a}}_{\text{mle}} - \mathbb{E}[\hat{\mathbf{a}}_{\text{mle}}])^T].$$

6. **Confidence interval construction:** Based on the mean and covariance matrix from the empirical distribution, 95% confidence intervals for each parameter are constructed as follows:

$$\hat{\zeta}_{\text{mle}} \pm 1.96 \cdot \text{SE}(\hat{\zeta}_{\text{mle}}),$$

$$\hat{\eta}_{\text{mle}} \pm 1.96 \cdot \text{SE}(\hat{\eta}_{\text{mle}}),$$

$$\hat{k}_{\text{mle}} \pm 1.96 \cdot \text{SE}(\hat{k}_{\text{mle}}).$$

Through this simulation study, the behavior of MLEs for GEMBD parameters is thoroughly evaluated, providing insights into their statistical properties and convergence patterns as sample size increases.

## 7. Simulation Procedure and Results

In this section, we assume the true parameter values to be $\zeta = 2$, $\eta = 1.5$, and $k = 1$ to carry out the simulation procedure outlined in Section 6.

### 7.1. Bias and mean squared error (MSE) analysis

We generate samples of sizes $n = 10, 11, \ldots, 200$ and repeat each process 100 times. For each parameter, the bias and MSE are computed, and the trace plots are shown in Figure 5.

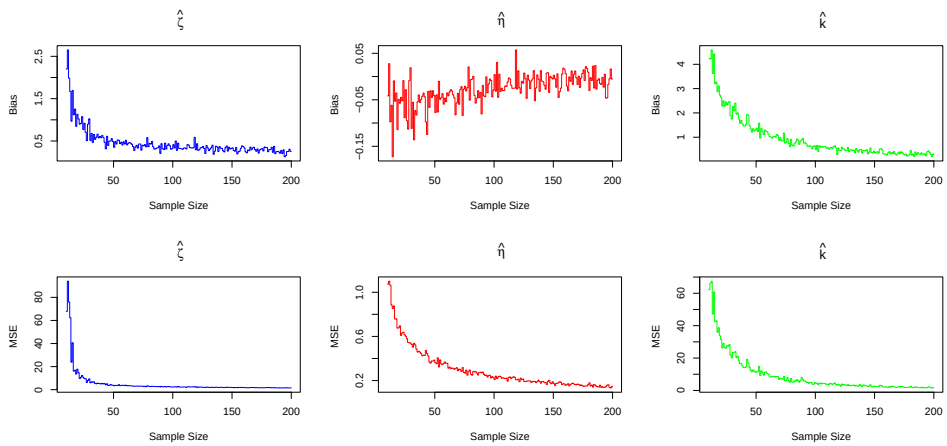



**Figure 5** Trace plots of bias and MSE for the MLEs

From Figure 5, we observe that the estimates for $\zeta$ and $k$ exhibit a positive bias, while $\eta$ is negatively biased. As the sample size increases, the bias in all parameters decreases and tends to approaches zero. The MSE follows a similar trend, where the values decrease as the sample size increases, eventually converging to zero.

### 7.2. Examination of asymptotic normality for MLEs

Samples of size $n = 20, 25, 30, \ldots, 850$ are generated, with the process repeated 100 times, producing 100 sets of parameter estimates for each sample size. For each set, the Anderson-Darling (AD) test statistic and corresponding p-values are calculated and the results are displayed as trace plots in Figure 6.

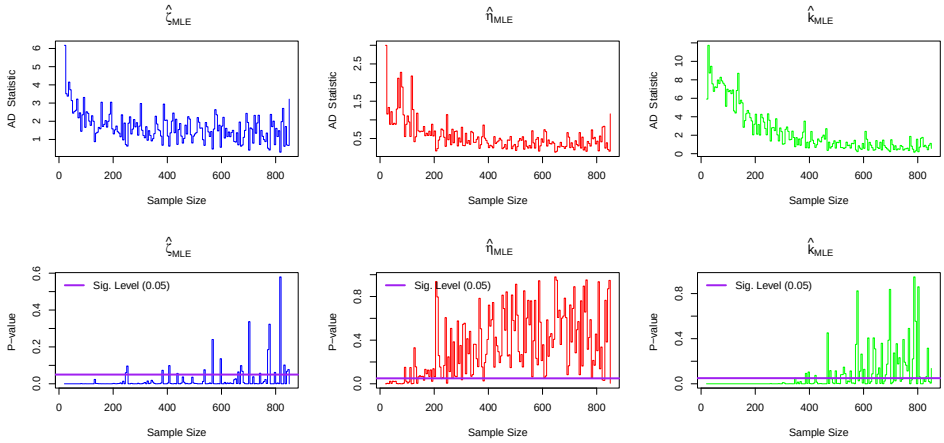



**Figure 6** Asymptotic normality behavior of MLEs for GEMBD parameters based on Anderson-Darling test

Figure 6 shows that as the sample size increases, the AD-test statistic steadily decreases and approaches zero for all three parameters. The corresponding p-values start crossing the 0.05 significance level with increasing frequency as the sample size grows. The MLEs for $\eta$ shows quick convergence to normality as compared to MLEs of $k$, while the MLE for $\zeta$ approaches normality more gradually.

The trace plots in Figure 6 confirm that the MLEs exhibit asymptotic normality as the sample size increases. A larger sample of size 600 was generated and repeated 100 times to obtain the asymptotic variance-covariance matrix and Fisher information matrix, resulting in an empirical sampling distribution for $\hat{\zeta}, \hat{\eta}, \hat{k}$. Based on this distribution, we calculated the variance-covariance matrix, Fisher information matrix, and 95% confidence intervals, as presented in Table 3. Additionally, a pair-plot of the estimates is shown in Figure 7.

**Table 3** Fisher Information and Covariance Matrix of Parameters of GEMBD

| | $\mathbf{a}_{\text{mle}}$ | $I^{-1}(\mathbf{a}_{\text{mle}})$ | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Parameters | | $\hat{\zeta}_{mle}$ | $\hat{\eta}_{mle}$ | $\hat{k}_{mle}$ | $\hat{\zeta}_{mle}$ | $\hat{\eta}_{mle}$ | $\hat{k}_{mle}$ | 95% Asymptotic C.I. |
| $\hat{\zeta}_{mle}$ | (2.1686) | 85.3476 | -435.9174 | 80.2725 | 1.2044 | 0.3110 | -0.8087 | $\hat{\zeta}_{mle}(0.0176, 4.3195)$ |
| $\hat{\eta}_{mle}$ | (1.4948) | -435.9174 | 2426.0433 | -332.5961 | 0.3110 | 0.0891 | -0.2425 | $\hat{\eta}_{mle}(0.9098, 2.0798)$ |
| $\hat{k}_{mle}$ | (1.1498) | 80.2725 | -332.5961 | 125.0891 | -0.8087 | -0.2425 | 0.7434 | $\hat{k}_{mle}(-0.5400, 2.8397)$ |

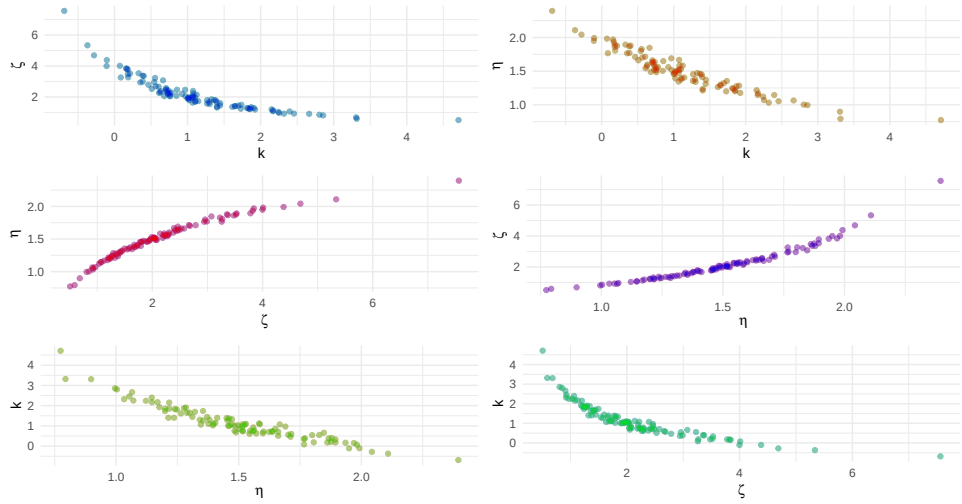



**Figure 7** Pair-plot showing correlations between MLEs of GEMBD parameters.

From Table 3, it can be seen that a higher Fisher information value results in a lower variance in the corresponding parameter estimates. Additionally, both Table 3 and Figure 7 show that $\hat{k}$ is negatively correlated with both $\hat{\zeta}$ and $\hat{\eta}$, whereas $\hat{\zeta}$ and $\hat{\eta}$ exhibit a positive correlation.

## 8. Applications

This section illustrates the practical utility and adaptability of the GEMBD model by analyzing two real-world datasets. The objective is to assess the model's versatility and compatibility in comparison to its sub-models and its base model. To accomplish this, we apply maximum likelihood estimation for parameter fitting and employ a variety of model selection criteria, including AIC, BIC, AICC, and HQIC. Generally, a model that achieves lower values for these criteria is considered superior.

**Dataset I:** This dataset from Faraway (2025) comprises clinical data from a study of 97 men with prostate cancer, each scheduled for radical prostatectomy, and represents the logarithm of prostate weight. Reported originally by Andrews and Herzberg (1985) (reprinted in 2012).

**Dataset II:** This dataset from Ahad et al. (2024) represents tensile strength measurements (in GPA) from 69 carbon fibers tested under tension at gauge lengths of 20mm. These measurements were originally reported by Bader and Priest (1982).

Tables 4 and 5 shows the maximum likelihood estimates and different information measure for dataset I and II respectively and Figure 8 shows the plots of estimated density functions of GEMBD and its submodels. Furthermore, the estimated CDF, Quantile-Quantile(QQ), KaplanMeier survival plot for dataset I and II are shown in Figure 9.

The results in Tables 4 and 5, clearly demonstrate that the proposed model consistently yields the smallest values across the model selection tools. Therefore, we conclude that the GEMBD distribution proves to be best model than its sub-models and its base model. Also the results are validated by Figure 8.

**Table 4** Estimates with corresponding standard error (provided in parenthesis) and Performance measure of the distribution for dataset I

| Model | Parameters | | | Information criterion tools | | | | |
|---|---|---|---|---|---|---|---|---|
| | $\zeta$ | $\eta$ | $k$ | $-2l$ | AIC | BIC | AICC | HQIC |
| GEMBD | 2.0325 (2.4109) | 1.5000 (0.5840) | 7.3598 (4.5972) | 139.349 | 145.349 | 153.073 | 145.607 | 148.472 |
| GMBD | 34.4774 (10.2415) | 2.9416 (0.2018) | - | 164.202 | 168.202 | 173.351 | 168.329 | 170.284 |
| WD | 6.3082 (0.3990 ) | 3.8730 (0.0662) | - | 168.603 | 172.603 | 177.753 | 172.731 | 174.685 |
| ABMD | 0.3680 (0.0236) | - | - | 228.202 | 230.202 | 232.777 | 230.244 | 231.243 |
| LBMD | 1.8430 (0.0662) | - | - | 247.992 | 249.992 | 252.567 | 250.034 | 251.033 |
| LBWRD | 10.1897 (760.1883) | 1.5000 ( 111.9056) | - | 247.992 | 251.992 | 257.141 | 252.120 | 254.074 |
| MBD | 2.1281 (0.0882) | - | - | 275.069 | 277.069 | 279.644 | 277.111 | 278.110 |
| SGD | 4.1097 (0.1935) | - | - | 306.988 | 308.988 | 311.562 | 309.030 | 310.029 |
| RD | 2.6064 (0.1323) | - | - | 316.050 | 318.050 | 320.624 | 318.092 | 319.091 |
| Chi | - | - | 4.5765 (0.2748) | 368.971 | 370.971 | 373.545 | 371.013 | 372.012 |
| HND | 3.6860 (0.2646) | - | - | 393.882 | 395.882 | 398.457 | 395.924 | 396.923 |
| ED | 3.6527 (0.3709) | - | - | 445.320 | 447.320 | 449.895 | 447.362 | 448.361 |

**Table 5** Estimates with corresponding standard error (provided in parenthesis) and Performance measure of the distribution for dataset II

| Model | Parameters | | | Information criterion tools | | | | |
|---|---|---|---|---|---|---|---|---|
| | $\zeta$ | $\eta$ | $k$ | $-2l$ | AIC | BIC | AICC | HQIC |
| GEMBD | 0.1978 (0.0881) | 0.4330 (0.1343) | 13.5950 (4.5972) | 113.607 | 119.607 | 126.036 | 120.014 | 122.136 |
| GMBD | 9.4473 (2.8679) | 2.2655 (0.2264) | - | 122.225 | 126.225 | 130.511 | 126.425 | 127.911 |
| WD | 5.0494 (0.4557) | 3.3147 (0.0878) | - | 123.914 | 127.914 | 132.200 | 128.114 | 129.600 |
| ABMD | 0.5134 (0.0409 ) | - | - | 138.222 | 140.222 | 142.366 | 140.288 | 141.065 |
| LBMD | 1.5603 (0.0695) | - | - | 148.330 | 150.330 | 152.473 | 150.395 | 151.172 |
| LBWRD | 3.1851 (126.0308) | 0.6541 (25.8823) | - | 148.330 | 152.330 | 156.616 | 152.530 | 154.015 |
| MBD | 1.8017 (0.0927) | - | - | 163.170 | 165.170 | 167.313 | 165.235 | 166.013 |
| RD | 2.2067 (0.1390) | - | - | 187.040 | 189.040 | 191.183 | 189.105 | 189.883 |
| SGD | 3.4860 (0.2187) | - | - | 190.756 | 192.756 | 194.899 | 192.822 | 193.599 |
| Chi | - | - | 3.9465 (0.3111) | 228.966 | 230.966 | 233.109 | 231.031 | 231.809 |
| HND | 3.1207 (0.2780) | - | - | 234.845 | 236.845 | 238.988 | 236.910 | 237.688 |
| ED | 3.0593 (0.3854) | - | - | 266.892 | 268.892 | 271.035 | 268.957 | 269.734 |

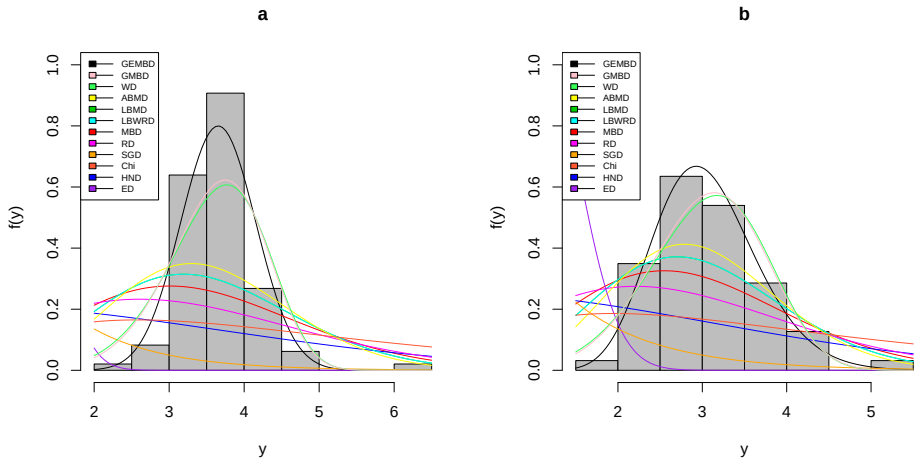



**Figure 8** (a) Plots of estimated PDF of GEMBD, its sub-models and base model for dataset I. (b) plots of estimated PDF of GEMBD, its sub-models and base model for dataset II

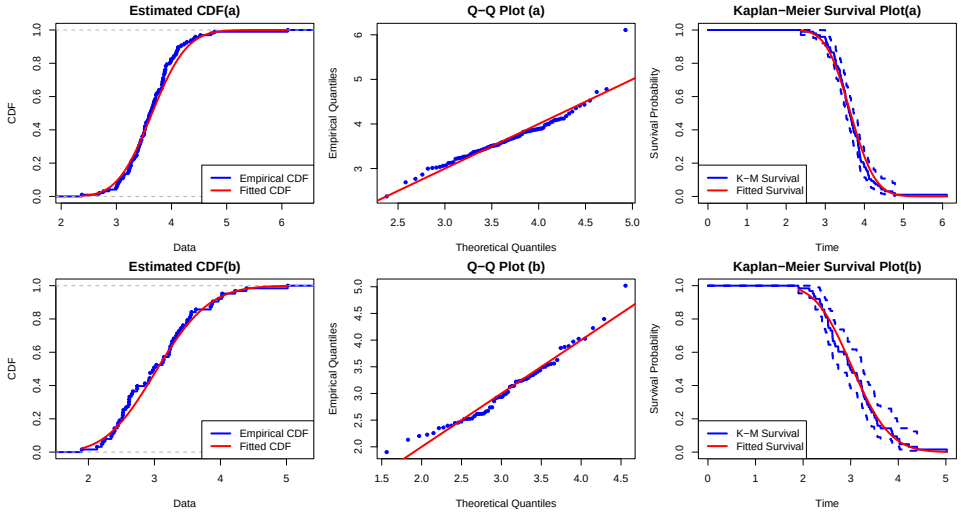



**Figure 9** (a) Estimated CDF , Q-Q plot , KaplanMeier survival plot for dataset I (b) Estimated CDF , Q-Q plot , KaplanMeier survival plot for dataset II

## 9. Conclusion

This manuscript introduces a new generalized model called the Generalized equilibrium Maxwell-Boltzmann distribution (GEMBD), developed by using the higher order weighted equilibrium technique. The study explores various statistical characteristics of this model, including aging behavior and stochastic ordering relations. We derive explicit formulas for crucial statistical properties such as moments, conditional moments, incomplete moments, mean residual life function, mean waiting time, Renyi entropy and densities of order statistics. Parameter estimation for the GEMBD is performed using maximum likelihood estimation techniques. Simulation studies reveal that as sample size grows, the MLEs converge asymptotically to a normal distribution, with observable reductions in bias and mean squared error (MSE). A positive correlation is observed between $\hat{\zeta}_{\text{MLE}}$ and $\hat{\eta}_{\text{MLE}}$,

while $k$ is negatively correlated with both $\hat{\zeta}_{\text{MLE}}$ and $\hat{\eta}_{\text{MLE}}$. Additionally, parameter $\hat{\eta}_{\text{MLE}}$ shows a slight negative bias in estimation, whereas parameters $\hat{\zeta}_{\text{MLE}}$ and $\hat{k}_{\text{MLE}}$ demonstrate positive biases. These biases decrease towards zero with larger sample sizes, indicating enhanced accuracy and reliability in parameter estimation as data availability increases. We also demonstrate real-life applications of the proposed model, highlighting its superior fit compared to competing models across two different real-life datasets. Furthermore, our analysis reveals that the Generalized equilibrium Maxewell Boltzmann distribution performs better compared to its sub-models and base model across all two datasets. Overall, the proposed GEMBD is a flexible and robust model that demonstrates strong fitting capability together with rich analytical properties. Nevertheless, the inclusion of additional parameters may increase computational complexity and slightly reduce interpretability when compared to simpler models. Also, the model can be employed for physical phenomena, however, the present study does not investigate the correspondence between model parameters and underlying physical processes. Future research may focus on developing physical interpretations of the model parameters, simplifying estimation procedures, extending the model to censored or dependent data and exploring broader applications in reliability engineering, survival analysis, and related fields.

**Acknowledgement**

The authors sincerely thank the referees for their valuable comments and constructive suggestions, which helped improve the manuscript.

**References**

Ahad N, Ahmad SP, Reshi JA. Bayesian analysis of extended Maxwell-Boltzmann distribution using simulated and real-life data sets. Reliab Theory Appl. 2024; 19(3): 188-201.

Ahad N, Ahmad SP, Reshi JA. A New Weighted-Generator Family of Statistical Mechanics Models With Simulation and Analysis to Light Intensity and Tensile Strength of Carbon Fiber Data. Rev Invest Opera. 2026; 47: 50-66.

Al-Khadim AK, Hussein AN. New proposed length biased weighted exponential and Rayleigh distribution with application. Math Theory Model. 2014; 4(7): 137-152.

Anas M, Dar AA, Ahmed A, Arshad M. K th-order equilibrium Weibull distribution: properties, simulation, and its applications. Commun Stat Simulat. 2024(12); 6237-6259.

Andrews DF, Herzberg AM. Data: a collection of problems from many fields for the student and research worker. 1985. (Reprinted in 2012)

Andrews DF, Herzberg AM. Data: a collection of problems from many fields for the student and research worker. Springer Sci Bus Media; 2012.

Bader MG, Priest AM. Statistical aspects of fibre and bundle strength in hybrid composites. Prog Sci Eng Compos. 1982; 1129-1136.

Bhat AA, Ahmad SP. On mixture of Maxwell and Weighted Maxwell distributions: properties, estimation and applications. Pak J Stat. 2022; 38(3): 359-376.

Faraway J. Package faraway. 2025.

Harkness W, Shantaram R. Convergence of a sequence of transformations of distribution functions. Pac J Math. 1969; 31(2): 403-415.

Helmert FR. ber die Berechnung des wahrscheinlichen Fehlers aus einer endlichen Anzahl wahrer Beobachtungsfehler. Z Math U Physik. 1875; 20: 300-303.

Helmert FR. Die genauigkeit der formel von peters zur berechnung des wahrscheinlichen beobachtungsfehlers director beobachtungen gleicher genauigkeit. Astron Nach. 1876; 88(8): 113.

Johnson NL, Kotz S, Balakrishnan N. Continuous univariate distributions. Vol. 2. John Wiley & Sons; 1995.

Leone FC, Nelson LS, Nottingham RB. The folded normal distribution. Technometrics. 1961; 3(4): 543-550.

Maxwell JC. On the dynamical theory of gases. Lond Edinb Dubl Philos Mag J Sci. 1868; 35(235): 129-145.
Modi K, Gill V. Length-biased weighted Maxwell distribution. Pak J Stat Oper Res. 2015; 465-472.
Nair NU, Preeth M. Multivariate equilibrium distributions of order n. Stat Probab Lett. 2008; 78(18): 3312-3320.
Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. Lond Edinb Dubl Philos Mag J Sci. 1900; 50(302): 157-175.
Rao CR. On discrete distributions arising out of methods of ascertainment. Sankhy: Ind J Stat. 1965; 311-324.
Rayleigh L. On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. Lond Edinb Dubl Philos Mag J Sci. 1880; 10(60): 73-78.
Reshi JA, Para BA, Bhat SA. Parameter Estimation of Weighted Maxwell-Boltzmann Distribution Using Simulated and Real Life Data Sets. Adapt Filter - Recent Adv Pract Implement. IntechOpen; 2021.
Saghir A, Khadim A, Lin Z. The Maxwell length-biased distribution: Properties and estimation. J Stat Theory Pract. 2017; 11(1): 26-40.
Shanker R, Zakaria AFM, Prodhani HR, Ray M. Weighted Uma distribution with properties and applications. J XIDIAN Univ. 2023; 17(10): 830-846.
Sharma VK, Dey S, Singh SK, Manzoor U. On length and area-biased Maxwell distributions. Commun Stat-Simul Comput. 2018; 47(5): 1506-1528.
Voda VGh. A modified Weibull hazard rate as generator of a generalized Maxwell distribution. Math Rep. 2009; 11(2): 171-179.
Weibull W. A statistical theory of the strength of material. Ingeniors Veten Acad Handl. 1939; 151(1): 1-45.
Willmot GE, Drekic S, Cai J. Equilibrium compound distributions and stop-loss moments. Scand Actuar J. 2005; 2005(1): 6-24.
Yu J, Kim SB, Bai J, Han S. Comparative Study on Exponentially Weighted Moving Average Approaches for the Self-Starting Forecasting. Appl Sci. 2020; 10: 7351.

## Appendix

**Theorem 1** *Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $H = [a, d]$ be an interval for some $a < d$ (where $a$ and $d$ may be extended to $-\infty$ or $\infty$). Let $Y$ be a continuous random variable with distribution function $F$, and let $b_1$ and $b_2$ be two real-valued functions defined on $H$ such that:*

$$E[b_1(Y)|_{Y>y}] = E[(b_2(y))|_{Y>y}]A(y), y \in H$$

*is defined in terms of some real function $A$. Assume that:*

1. *$b_1(Y)$ and $b_2(Y)$ are differentiable, and $F$ is twice continuously differentiable and strictly monotone on $H$.*
2. *The equation $A(Y)b_1(Y) = b_2(Y)$ has no real solution within the interior of $H$.*

*Under these conditions, $F$ is uniquely determined by the functions $b_1$, $b_2$, and $A$. In particular:*

$$F(y) = C \int_a^y \left| \frac{A'(u)}{A(u)b_1(u) - b_2(u)} \right| exp(S(u))\, du$$

*where $S$ is a solution to the differential equation:*

$$s'(y) = \frac{A'(y)b_2(y)}{A(y)b_1(y) - b_2(y)}$$

*and $C$ is a normalization constant such that $\int_H F(y)\, dy = 1$.*