The odd Chen generator of distributions: properties and estimation methods with applications in medicine and engineering

M El-Morshedy\textsuperscript{1,2*}, MS Eliwa\textsuperscript{2} and AZ Affy\textsuperscript{3}

\textsuperscript{1} Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia.
\textsuperscript{2} Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt.
\textsuperscript{3} Department of Statistics, Mathematics and Insurance, Benha University, Benha, Egypt.

Submitted: 13 February 2019; Revised: 16 July 2019; Accepted: 22 November 2019

Abstract: This paper introduces a new univariate flexible generator of distributions called the odd Chen-G family, and some of its statistical properties are derived. Two special models of the proposed generator are provided. The model parameters are estimated using six estimation methods, namely, maximum likelihood estimators, least squares estimators, weighted least squares estimators, maximum product of spacings estimators, Cramèr-von Mises estimators and percentile based estimators. Further, simulations are performed to compare their performances for both small and large samples. Finally, two real datasets are used to illustrate the flexibility of the special models of the proposed family.

Keywords: Chen distribution, estimation methods, hazard rate function, maximum likelihood estimators, simulation, weighted least squares estimators.

INTRODUCTION

Numerous classical distributions have been extensively used over the past decades for modelling data in several areas such as medical sciences, life testing problems, biological studies, demography, engineering, actuarial, environmental, and economics. However, in many applied areas such as insurance, survival and reliability theory, there is a clear need for extended forms of these distributions, because, in many practical situations, classical distributions do not provide adequate fits to real data. Therefore, there has been an increased interest in developing more flexible distributions; for example, El-Gohary et al. (2015a, 2015b), El-Bassiouny et al. (2016, 2017), El-Morshedy et al. (2017), Eliwa et al. (2018) and Jehhan et al. (2018) among others.

Recently, several families employing one or more parameters to generate new distributions have been proposed in the literature; for example, Alzaatreh et al. (2013), Amini et al. (2018), Haghbin et al. (2016), El-Morshedy and Eliwa (2019) and Cordeiro et al. (2019), Alizadeh et al. (2020), among others.

Chen (2000) proposed a new two-parameter distribution with bathtub shaped or increasing failure rate function called the Chen (Ch) distribution with positive parameters $\alpha$ and $\beta$. The cumulative distribution function (CDF) and probability density function (PDF) of the Ch distribution are

\begin{equation}
\Pi(t; \alpha, \beta) = 1 - \exp\left(-\alpha \left[ \exp\left(t^\beta\right) - 1 \right] \right), \quad t > 0
\end{equation}

and

\begin{equation}
\pi(t; \alpha, \beta) = \alpha \beta t^{\beta-1} \exp\left(t^\beta\right) \exp\left(-\alpha \left[ \exp\left(t^\beta\right) - 1 \right] \right), \quad t > 0.
\end{equation}

Several extensions of the Ch distribution are discussed in the statistical literature; see for example Chaubey and Zhang (2015), Dey et al. (2017) and Shuaib et al. (2018),

\footnote{Corresponding author (m.elmorshedy@psau.edu.sa; http://orcid.org/0000-0002-7585-5519)}

This article is published under the Creative Commons CC-BY-ND License (http://creativecommons.org/licenses/by-nd/4.0/). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in any way.
among others. Due to the importance of the Ch distribution, we introduce a new class called the odd Chen-G (OCh-G) family of distributions.

**METHODOLOGY**

Assume $T$ is the lifetime of a system following the Ch CDF in equation (1). If the random variable $X$ represents the odds ratio, the risk that this system will not be working at time $x$ is given by $G(x; \eta)/[1 - G(x; \eta)]$, where $G(x; \eta)$ represents the CDF of a baseline model with parameter vector $\eta$ (1 $\times$ $k$).

$$F(x) = P(X \leq x) = \Pi \left( \frac{G(x; \eta)}{1 - G(x; \eta)} \right),$$

where the odds ratio $G(x; \eta)/[1 - G(x; \eta)]$ satisfies the following conditions:

1. $G(x; \eta)/[1 - G(x; \eta)] \in [a^*, b^*]$ for $0 < a^* < b^* < \infty$.
2. $G(x; \eta)/[1 - G(x; \eta)]$ is differentiable and monotonically non-decreasing.
3. $G(x; \eta)/[1 - G(x; \eta)] \to a^*$ as $x \to 0$ but $G(x; \eta)/[1 - G(x; \eta)] \to b^*$ as $x \to \infty$.

Let $g(x; \eta), G(x; \eta),$ and $C(x; \eta)$, respectively denote the PDF, CDF and RF of a baseline model with parameter vector $\eta$ (1 $\times$ $k$). Then, the CDF of the OCh-G family is given by

$$F(x; \alpha, \beta, \eta) = \int_0^{G(x; \eta)} \pi(t)dt \exp \left[-\alpha \left\{ \exp \left( \frac{G(x; \eta)}{1 - G(x; \eta)} \right)^\beta - 1 \right\} \right].$$

The PDF of the OCh-G family is given as follows

$$f(x; \alpha, \beta, \eta) = \frac{\alpha \beta g(x; \eta)G(x; \eta)^{\beta - 1}}{[1 - G(x; \eta)]^{\beta + 1}} \exp \left[ \frac{G(x; \eta)}{1 - G(x; \eta)} \right]^{\gamma} \exp \left(-\alpha \left\{ \exp \left( \frac{G(x; \eta)}{1 - G(x; \eta)} \right)^\beta - 1 \right\} \right).$$

Using the power series for the exponential function and the generalised binomial expansion, equations (3) and (4) can be represented as an infinite mixture of exponential-G (Exp-G) density functions as follows

$$F(x; \alpha, \beta, \eta) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} C_{i,j,k,m}(\alpha, \beta)h_{\beta k + m}(x),$$

and

$$f(x; \alpha, \beta, \eta) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} C_{i,j,k,m}(\alpha, \beta)h_{\beta k + m}(x),$$

respectively, where $h_{\beta k + m}(x) = G(x; \eta)^{\beta k + m}$ is the $\text{Exp} - G$ family with power parameter $(\beta k + m), h_{\beta k + m}(x) = (\beta k + m)g(x; \eta)G(x; \eta)^{\beta k + m - 1}$, and

$$C_{i,j,k,m}(\alpha, \beta) = \frac{(-1)^{i+j} \alpha^i (i-j)^j \Gamma(\beta k + m)}{i! \Gamma(\beta k)},$$

where $(\beta k) \neq 0, -1, -2, -3, ...$ and $(\beta k + m) \neq 0, -1, -2, -3, ...$. The reliability function (RF) of the OCh-G class is

$$R(x; \alpha, \beta, \eta) = \exp \left(-\alpha \left\{ \exp \left( \frac{G(x; \eta)}{1 - G(x; \eta)} \right)^\beta - 1 \right\} \right); \quad x > 0.$$

The hazard rate function (HRF) is

$$h(x; \alpha, \beta, \eta) = \frac{\alpha \beta g(x; \eta)G(x; \eta)^{\beta - 1}}{[1 - G(x; \eta)]^{\beta + 1}} \exp \left( \frac{G(x; \eta)}{1 - G(x; \eta)} \right); \quad x > 0.$$

There are following motivations for introducing the OCh-G family:

i. To construct heavy-tailed distributions for modelling real data.
ii. To define special models with all types of the HRF.
iii. To make the kurtosis more flexible compared to the baseline model.
iv. To provide consistently better fits than other generated models under the same baseline distribution.
v. To generate distributions with symmetric, left-skewed and right-skewed shaped.
vi. To study which is the best method to estimate the model parameters.

Furthermore, we are also motivated to illustrate how different estimators of the special sub models of the OCh-G class perform for various sample sizes and various parameter combinations and to develop a guideline for determining the best method of estimation, which is important for the applied statisticians.

Statistical properties

Asymptotics

Let \( \alpha = \inf \{x | G(x; \eta) > 0\} \), then the asymptotics of the CDF, PDF and HRF when \( x \to \alpha \) are, respectively, given by

\[
F(x; \alpha, \beta, \eta) \sim \alpha G(x; \eta)^\beta
\]

\[
f(x; \alpha, \beta, \eta) \sim \alpha \beta g(x; \eta) G(x; \eta)^{\beta-1}
\]
and

\[
h(x; \alpha, \beta, \eta) \sim \alpha \beta g(x; \eta) G(x; \eta)^{\beta-1}.
\]

Further, the asymptotics of the CDF, PDF and HRF when \( x \to \infty \) are, respectively, given by

\[
1 - F(x; \alpha, \beta, \eta) \sim e^{-\alpha e^{(1 - G(x; \eta))^{-\beta}}},
\]

\[
f(x; \alpha, \beta, \eta) \sim \alpha \beta g(x; \eta) [1 - G(x; \eta)]^{\beta-1} e^{(1 - G(x; \eta))^{-\beta}}
\]

\[
h(x; \alpha, \beta, \eta) \sim \alpha \beta g(x; \eta) [1 - G(x; \eta)]^{\beta-1} e^{(1 - G(x; \eta))^{-\beta}}.
\]

Quantile function

Assume \( X \sim \text{OCh-G} \) family, for any \( u^* \in (0,1) \), the \( u^* \)th quantile function, say \( Q(u^*) \), of \( X \) is the solution of \( F(Q(u^*)) - u^* = 0 \); \( Q(u^*) > 0 \), then

\[
Q(u^*) = G^{-1}\left(\frac{1}{\frac{1}{\log \left(1 - \frac{1}{u^*} \log \left(1 - u^*\right)\right)}^{-1/\beta}}\right),
\]

where \( G^{-1} \) represents the baseline quantile function. Setting \( u^* = 0.5 \), we get the median of \( X \).

Moments, skewness, kurtosis and mean deviations

If \( X \sim \text{OCh-G} \) family, the \( r \)th moment of \( X \) is given by

\[
\mu'_r = E(X^r) = \int_0^\infty x^r f(x; \alpha, \beta, \eta) dx
\]

\[
= \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \sum_{k,m=0}^{\infty} Y_{i,k,j,m}(\alpha, \beta) \int_0^\infty x^r h_{\beta+k+m}(x) dx
\]

\[
= \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \sum_{k,m=0}^{\infty} Y_{i,k,j,m}(\alpha, \beta) E(Y^r_{\beta+k+m}),
\]

... (10)

where \( Y^r_{\beta+k+m} \sim \text{Exp} - G \) with power parameter \( \beta k + m \). Setting \( r = 1 \) in equation (10), we obtain the mean of \( X \). Moreover, we can derive an important measure in survival analysis called the mean time to failure (MTTF) where \( \text{MTTF} = \mu'_1 \). This measure can be used in order to design and manufacture a maintainable system.

On the other hand, the skewness and the kurtosis can be calculated, respectively, as follows:

\[
\delta_1 = \left( \mu_3 - 3\mu'_1 \mu'_2 + 2\mu'_1^3 \right) / \left( \mu'_2 - \mu'_1^2 \right)^{3/2}
\]

... (11)

\[
\delta_2 = \left( \mu'_4 - 4\mu'_2 \mu'_2 + 6\mu'_1 \mu'_2 - 3\mu'_1^2 \right) / \left( \mu'_2 - \mu'_1^2 \right)^2.
\]

... (12)

Further, the incomplete moments play an important role for measuring inequality. For example, the first incomplete moment can be used to obtain the formulas of Lorenz and Bonferroni curves. The \( q \)th incomplete moment of \( X \) can be expressed as follows:

\[
\Phi_q(t) = \int_0^t x^q f(x; \alpha, \beta, \eta) dx
\]

\[
= \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \sum_{k,m=0}^{\infty} Y_{i,k,j,m}(\alpha, \beta) \int_0^t x^q h_{\beta+k+m}(x) dx
\]

\[
= \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \sum_{k,m=0}^{\infty} Y_{i,k,j,m}(\alpha, \beta) \bar{\Phi}_q^*(t),
\]

... (13)
where \( \Phi^*_q(t) = \int_0^t x^q h_{2k+m}(x)dx \). On the other hand, the mean deviations about the mean and the median can be represented as

\[
\lambda_1 = 2\mu_1 P(\mu_1) - 2\Phi(t_1)(\mu_1)
\]

and

\[
\lambda_2 = \mu_1 - 2\Phi(Q(0.5)),
\]

respectively.

**Bonferroni and Lorenz curves**

Bonferroni (1930) presented the Bonferroni and Lorenz curves. These curves have applications in reliability, medicine, demography, economics and insurance. If \( X \sim \text{OCh-G} \) family, then the Bonferroni curve is given by

\[
B(p) = \frac{1}{\mu^*} \int_0^{G^{-1}(p)} x f(x; \alpha, \beta, \eta)dx
\]

where \( B_{2k+m}(p) = \frac{1}{\mu^*} \int_0^{G^{-1}(p)} x h_{2k+m}(x)dx \) is the Bonferroni curve of the Exp-G family with power parameter \((2k + m)\) and \(\mu^*\) denotes the average. The Lorenz curve can be expressed as

\[
L(p) = \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \sum_{k,m=0}^{\infty} Y_{i,j,k,m}(\alpha, \beta) B_{2k+m}(p),
\]

where \( L_{2k+m}(p) = \frac{1}{\mu^*} \int_0^{G^{-1}(p)} x h_{2k+m}(x)dx \) is the Lorenz curve of the Exp-G family with power parameter \((2k + m)\).

**Moments of the residual and past lifetimes**

For describing different maintenance strategies, we must calculate two important times, namely, the mean residual lifetime (MRL) and the mean past lifetime (MPL). The \( n \)th moment of the residual lifetime (RL) is given as

\[
M_{RL}^{(n)}(t) = \mathbb{E}((T-t)^n | T > t) ; n = 1, 2, 3, ...
\]

Therefore, if \( T \sim \text{OCh-G} \) family, then

\[
M_{RL}^{(n)}(t) = \frac{1}{R(t; \alpha, \beta, \eta)} \int_t^{\infty} (x-t)^n f(x; \alpha, \beta, \eta)dx
\]

\[
= \frac{1}{R(t; \alpha, \beta, \eta)} \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \sum_{k,m=0}^{\infty} Y_{i,j,k,m}(\alpha, \beta) \int_t^{\infty} (x-t)^n h_{2k+m}(x)dx
\]

\[
= \frac{1}{R(t; \alpha, \beta, \eta)} \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \sum_{k,m=0}^{\infty} \sum_{l=0}^{n} (-1)^l \binom{n}{l} \left( \frac{n}{l} \right) Y_{i,j,k,m}(\alpha, \beta) M_{RL}^{(n-l)}(t),
\]

where \( M_{RL}^{(n)}(t) = \int_0^{\infty} x^{n-1} h_{2k+m}(x)dx \). Setting \( n = 1 \) in equation (16), we get the MRL.

The \( n \)th moment of the past lifetime (PL) (also known as the \( n \)th moment of the waiting time) can be expressed as

\[
M_{PL}^{(n)}(t) = \mathbb{E}((t-T)^n | T \leq t)
\]

\[
= \frac{1}{F(t; \alpha, \beta, \eta)} \int_0^{t} (t-x)^n f(x; \alpha, \beta, \eta)dx
\]

\[
= \frac{1}{F(t; \alpha, \beta, \eta)} \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} \sum_{k,m=0}^{\infty} \sum_{l=0}^{n} (-1)^l \binom{n}{l} \left( \frac{n}{l} \right) Y_{i,j,k,m}(\alpha, \beta) M_{PL}^{(n-l)}(t),
\]

where \( M_{PL}^{(n)}(t) = \int_0^{\infty} x^{n-1} h_{2k+m}(x)dx \). Setting \( n = 1 \) in equation (17), we get the MPL.

**Special OCh-G models**

**The OCh-Fréchet (OChFr) distribution**

Consider the CDF of the Fréchet distribution with positive parameters \( a \) and \( b \) given by

\[
G(x; a, b) = e^{-(\frac{x}{b})^a}, \quad x > 0.
\]
Then, the CDF of the OChFr distribution is

\[ F(x; \alpha, \beta, a, b) = 1 - \exp \left( -\alpha \left[ \exp \left( \frac{a}{x} \right)^b - 1 \right]^{-\beta} - 1 \right); \quad x > 0. \]

Figure 1 shows some plots of the PDF and the HRF of the OChFr distribution for various values of the parameters.

Figure 1 shows that the HRF of the OChFr model can be decreasing, increasing or unimodal-shaped. The skewness and kurtosis of the OChFr distribution for some choices of \( a = 2, b = 2.5 \) and \( \beta \) as function of \( \alpha \) are displayed in Figure 2.

Figure 1: The PDF plots (left panel) and the HRF plots (right panel) of the OChFr distribution

Figure 2: The skewness plots (left panel) and the kurtosis plots (right panel) of the OChFr distribution
Figure 2 reveals that the shapes of the OChFr model have strong dependence on the values of $\alpha$ and $\beta$.

**The OCh-Weibull (OChW) distribution**

Consider the CDF of the Weibull distribution with positive parameters $a$ and $b$ given by $G(x; a, b) = 1 - e^{-(x)^{a}}$, $x > 0$.

Then, the CDF of the OChW distribution is

$$F(x; \alpha, \beta, a, b) = 1 - \exp \left[ -\alpha \left( \exp \left( \frac{x}{b} \right)^{a} - 1 \right)^{\beta} - 1 \right], \ x > 0.$$  ...(19)

Figure 3 shows the PDF and HRF plots of the OChW distribution for various values of the parameters. It can be seen that the HRF can be increasing or bathtub shaped. Further, the skewness and kurtosis of the OChW distribution for some choices of $a = 2$, $b = 0.5$ and $\beta$ as function of $\alpha$ are displayed in Figure 4.

It is clear from Figure 4, that the shapes of the OChW model have strong dependence on the values of $\alpha$ and $\beta$.

**Estimation methods**

In this section, six methods of estimation are used to estimate the unknown parameters of the OChFr and OChW models to illustrate how different estimators

---

*Figure 3: The PDF plots (left panel) and the HRF plots (right panel) of the OChW distribution*

*Figure 4: The skewness plots (left panel) and the kurtosis plots (right panel) of the OChW distribution*
of these distributions perform for various sample sizes and various parameter combinations and to develop a guideline for determining the best method of estimation, which is important for the applied statisticians.

These estimation methods are: the maximum likelihood estimators, least squares estimators, weighted least squares estimators, maximum product of spacings estimators, Cramér-von Mises estimators and percentile based estimators. Similar studies for other models have been proposed by many authors (Eliwa et al., 2018; Nassar et al., 2018; Cordeiro et al., 2019).

**Maximum likelihood estimators**

Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ from $F(x; \alpha, \beta, a, b)$. Then, the maximum likelihood estimators (MLEs) of the OChFr and OChW parameters $\hat{\alpha}_{MLE}$, $\hat{\beta}_{MLE}$, $\hat{a}_{MLE}$ and $\hat{b}_{MLE}$ can be obtained by maximizing

$$
L(\alpha, \beta, a, b) = \sum_{i=1}^{n} \log f(x_i; \alpha, \beta, a, b), \quad \text{(20)}
$$

with respect to $\alpha, \beta, a$ and $b$. Or equivalently, the MLEs follow by solving the non-linear equations defined by

$$
L_\alpha = \frac{n}{\alpha} - \sum_{i=1}^{n} \left[ \exp \left\{ \left[ \frac{G(x_i; a, b)}{1 - G(x_i; a, b)} \right]^{\beta} - 1 \right\} \right] = 0,
$$

$$
L_\beta = \frac{n}{\beta} + \sum_{i=1}^{n} \ln \left[ \frac{G(x_i; a, b)}{1 - G(x_i; a, b)} \right] - \sum_{i=1}^{n} \ln \left[ 1 - G(x_i; a, b) \right] + \sum_{i=1}^{n} \left[ \frac{G(x_i; a, b)}{1 - G(x_i; a, b)} \right]^{\beta} \ln \left[ \frac{G(x_i; a, b)}{1 - G(x_i; a, b)} \right]
$$

$$
- \alpha \sum_{i=1}^{n} \exp \left\{ \left[ \frac{G(x_i; a, b)}{1 - G(x_i; a, b)} \right]^{\beta} \right\} \left[ \frac{G(x_i; a, b)}{1 - G(x_i; a, b)} \right]^{\beta} \ln \left[ \frac{G(x_i; a, b)}{1 - G(x_i; a, b)} \right] = 0,
$$

$$
L_a = \sum_{i=1}^{n} \frac{[g(x_i; a, b)]_b}{g(x_i; a, b)} + (\beta - 1) \sum_{i=1}^{n} \frac{[G(x_i; a, b)]_b}{G(x_i; a, b)} + (\beta + 1) \sum_{i=1}^{n} \frac{[G(x_i; a, b)]_b}{1 - G(x_i; a, b)}
$$

where $L_* = \frac{\partial}{\partial A} L(\alpha, \beta, a, b)$ and $[A]^* = \frac{\partial}{\partial A} A$.

**Ordinary and weighted least-square estimators**

Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the order statistics of the random sample of size $n$ from $F(x; \alpha, \beta, a, b)$. The least square estimators (LSEs) of the OChFr and OChW parameters $\hat{\alpha}_{LS}$, $\hat{\beta}_{LS}$, $\hat{a}_{LS}$ and $\hat{b}_{LS}$ can be obtained by minimizing

$$
V(\alpha, \beta, a, b) = \sum_{i=1}^{n} \left[ F(x_{(i)}; \alpha, \beta, a, b) - \frac{1}{n+1} \right]^2
$$

with respect to $\alpha, \beta, a$ and $b$. Or equivalently, the LSEs follow by solving the non-linear equations defined by

$$
\sum_{i=1}^{n} \left[ F(x_{(i)}; \alpha, \beta, a, b) - \frac{1}{n+1} \right] \Delta_s(x_{(i)}; \alpha, \beta, a, b) = 0,
$$

$s = 1, 2, 3, 4,$

where

$$
\Delta_1(x_{(i)}; \alpha, \beta, a, b) = \frac{\partial}{\partial \alpha} F(x_{(i)}; \alpha, \beta, a, b)
$$

$$
\Delta_2(x_{(i)}; \alpha, \beta, a, b) = \frac{\partial}{\partial \beta} F(x_{(i)}; \alpha, \beta, a, b)
$$

$$
\Delta_3(x_{(i)}; \alpha, \beta, a, b) = \frac{\partial}{\partial a} F(x_{(i)}; \alpha, \beta, a, b)
$$

$$
\Delta_4(x_{(i)}; \alpha, \beta, a, b) = \frac{\partial}{\partial b} F(x_{(i)}; \alpha, \beta, a, b)
$$

\text{(21)}
Note that the solution of $\Delta_s$ for $s = 1, 2, 3, 4$ can be obtained numerically.

The weighted least squares estimators (WLSEs) $\hat{\alpha}_{WLS}, \hat{\beta}_{WLS}, \hat{\alpha}_{WLS}$ and $\hat{b}_{WLS}$ can be obtained by minimizing the following equation:

$$W(\alpha, \beta, a, b) = \sum_{i=1}^{n} \frac{(n+1)^2(n+2)}{i(n-i+1)} \left[ F(x(i)|\alpha, \beta, a, b) - \frac{i}{n+1} \right]^2.$$  

Further, the WLSEs can also be derived by solving the non-linear equations defined by

$$\Delta_s(x(i)|\alpha, \beta, a, b) = 0, \ s = 1, 2, 3, 4,$$

where $\Delta_1(\cdot|\alpha, \beta, a, b)$, $\Delta_2(\cdot|\alpha, \beta, a, b)$, $\Delta_3(\cdot|\alpha, \beta, a, b)$ and $\Delta_4(\cdot|\alpha, \beta, a, b)$ are provided in equation (21).

**Maximum product of spacings estimators**

The maximum product of spacings method is proposed as a good alternative to the MLE method. For $i = 1, 2, \ldots, n+1$, let

$$D_i(\alpha, \beta, a, b) = F(x(i)|\alpha, \beta, a, b) - F(x(i-1)|\alpha, \beta, a, b),$$

be the uniform spacings of a random sample from the OChFr distribution or OChW distribution, where $F(x(0)|\alpha, \beta, a, b) = 0$, $F(x(n+1)|\alpha, \beta, a, b) = 1$ and $\sum_{i=1}^{n+1} D_i(\alpha, \beta, a, b) = 1$. The maximum product of spacings estimators (MPSEs) for $\hat{\alpha}_{MPSE}, \hat{\beta}_{MPSE}, \hat{\alpha}_{MPSE}$ and $\hat{b}_{MPSE}$ can be obtained by maximizing the geometric mean of the spacings

$$G(\alpha, \beta, a, b) = \left[ \prod_{i=1}^{n+1} D_i(\alpha, \beta, a, b) \right]^\frac{1}{n+1},$$

with respect to $\alpha, \beta, a$ and $b$. Or by maximizing the logarithm of the geometric mean of sample spacings

$$H(\alpha, \beta, a, b) = \frac{1}{n+1} \sum_{i=1}^{n-1} \log D_i(\alpha, \beta, a, b).$$

The MPSEs can also be derived by solving the non-linear equations defined by

$$\frac{1}{n+1} \sum_{i=1}^{n+1} \frac{1}{D_i(\alpha, \beta, a, b)} \left[ \Delta_s(x(i)|\alpha, \beta, a, b) - \Delta_s(x(i-1)|\alpha, \beta, a, b) \right] = 0, \ s = 1, 2, 3, 4,$$

where $\Delta_1(\cdot|\alpha, \beta, a, b)$, $\Delta_2(\cdot|\alpha, \beta, a, b)$, $\Delta_3(\cdot|\alpha, \beta, a, b)$ and $\Delta_4(\cdot|\alpha, \beta, a, b)$ are defined in equation (21).

**Cramér-von Mises minimum distance estimators**

Cramér-von Mises estimators (CVMEs) are a type of minimum distance estimators and have less bias than the other minimum distance estimators. The CVMEs are obtained based on the difference between the estimates of the CDF and the empirical distribution function. The CVMEs of the OChFr and OChW parameters are obtained by minimizing

$$C(\alpha, \beta, a, b) = \frac{1}{12n} + \sum_{i=1}^{n} \left[ F(x(i)|\alpha, \beta, a, b) - \frac{2i-1}{2n} \right]^2,$$

with respect to $\alpha, \beta, a$ and $b$. Also, the CVMEs follow by solving the non-linear equations

$$\sum_{i=1}^{n} \left[ F(x(i)|\alpha, \beta, a, b) - \frac{2i-1}{2n} \right] \Delta_s(x(i)|\alpha, \beta, a, b) = 0, \ s = 1, 2, 3,$$

where $\Delta_1(\cdot|\alpha, \beta, a, b)$, $\Delta_2(\cdot|\alpha, \beta, a, b)$, $\Delta_3(\cdot|\alpha, \beta, a, b)$ and $\Delta_4(\cdot|\alpha, \beta, a, b)$ are defined in equation (21).

**Percentile based estimators**

Let $u_i = i/(n+1)$ be an unbiased estimator of $F(x(i)|\alpha, \beta, a, b)$. Hence, the percentile estimators (PCEs) of the OChFr and OChW parameters can be obtained by minimizing

$$P(\alpha, \beta, a, b) = \sum_{i=1}^{n} \left( x(i) - Q(u_i) \right)^2,$$

with respect to $\alpha, \beta, a$ and $b$, where $Q(u_i)$ is the quantile function of the OChFr and OChW distributions which are given, respectively, by using equation (9).
In this section, a simulation study is conducted to compare the performance of the MLEs, LSEs, WLSEs, MPSEs, CVMEs and PCEs of the unknown parameters for the OChFr and OChW distributions with respect to sample size \( n \): Mathcad V.15 is used herein to generate the samples as follows:

1. Generate 10000 samples of size \( n = 50, 150, 300, 500 \) from OChFr \((\alpha, \beta, a, b) = (0.5, 1.5, 2.0, 2.5)\) and OChW\((\alpha, \beta, a, b) = (2.5, 2.0, 1.5, 0.5)\), respectively.
2. Compute the MLEs, LSEs, WLSEs, MPSEs, CVMEs and PCEs for the 10,000 samples, say \( \theta_j \) for \( j = 1, 2, \ldots, 10,000 \).
3. Compute the average values of estimates (AEs) and mean-squared errors (MSEs).
The empirical results are given in Tables 1 and 2, respectively.

Regarding Tables 1 and 2, the following observations can be made:
1. The magnitude of bias always decreases to zero as $n \to \infty$.
2. The MSEs decrease when the sample size increases as expected under first-order asymptotic theory.
3. Depending on the MSEs, the MLE, LSE, WLSE, MPSE, CVME and PUCE methods perform quite well for estimating the OChFr and OChW parameters. However, we can consider the MLE, WLSE and CVME methods outperform LSE, MPSE and PCE methods. Therefore, the MLE, WLSE and CVME method are the best.

RESULTS AND DISCUSSION: DATA ANALYSIS

In this section, we illustrate the empirical importance of the OChFr and OChW distributions using two applications to real data.

The first dataset (I): represents the survival times, in weeks, of 33 patients suffering from acute Myelogeneous Leukaemia (Feigl & Zelen, 1965). For the dataset I, we shall compare the fits of the OChFr distribution with some competitive models listed in Table 3.

The second dataset (II): represents 40 observations of time-to-failure ($10^3$ h) of turbocharger of one type of engine (Xu et al., 2003). This dataset is used to compare the fits of the OChW model with some competitive models provided in Table 4.

The fitted distributions are compared using some criteria, namely, the maximized log likelihood ($-2L$), Akaike information criterion (AIC), Cramér-von Mises ($W^*$) statistic, Anderson-Darling ($A^*$) statistic, Kolmogorov-Smirnov (KS) statistic and its p value. Tables 5 and 7 list the MLEs with their corresponding standard errors (SEs) (in parentheses) for both datasets, respectively, whereas Tables 6 and 8 provide the values of goodness-of-fit measures for datasets I and II, respectively.

Table 3: The competitive models of the OChFr distribution for dataset I

| Distribution | Abbreviation | Author(s) |
|--------------|--------------|-----------|
| Weibull Fr   | $WF_r$       | Afify et al. (2016b) |
| Kumaraswamy Marshall-Olkin Fr | $KMOFr$ | Afify et al. (2016a) |
| Kumaraswamy Fr | $KF_r$ | Mead and Abd-Eltawab (2014) |
| Exponentiated Fr | $EF_r$ | Nadarajah and Kotz (2003) |
| Transmuted exponentiated Fr | $TEF_r$ | Elbatal et al. (2014) |
| Beta Fr | $BEF_r$ | Mead et al. (2017) |
| Fr | $F_r$ | Frechet (1924) |

Table 4: The competitive models of the OChW distribution for dataset II

| Distribution | Abbreviation | Author(s) |
|--------------|--------------|-----------|
| Odd log-logistic exponentiated W | $OLLEW$ | Afify et al. (2018) |
| Alpha logarithmic transformed W | $ALTW$ | Nassar et al. (2018) |
| Beta W | $BW$ | Lee et al. (2007) |
| Marshall-Olkin W | $MOW$ | Ghanty et al. (2005) |
| Transmuted complementary W geometric | $TCWG$ | Afify et al. (2014) |
| Lindley W | $LiW$ | Cordeiro et al. (2018) |
| W | $W$ | Weibull (1951) |
The values in Tables 6 and 8 show that the OChFr and OChW distributions have the lowest values of \(-2L\), AIC, \(W^*\), \(A^*\) and KS measures and then provide the best fits to both datasets. Furthermore, the p value test for the OChFr and OChW models have the largest value among all models. Hence, the OChFr and OChW distributions yield a better fit to those datasets than other distributions.

Figures 5 – 10 show the fitted PDFs, estimated CDFs and P-P plots of the tested distributions for both data sets.

### Table 5: The MLEs with their SEs for dataset I

| Model  | MLEs                      | SEs          |
|--------|---------------------------|--------------|
| OChFr  | \(\hat{\alpha} = 0.197,\) \(\hat{\beta} = 0.210,\) \(\hat{\delta} = 5.468,\) \(\hat{\gamma} = 1.490\) | \(0.099,\) \(0.092,\) \(2.803,\) \(0.607\) |
| WFr    | \(\hat{\alpha} = 2.622,\) \(\hat{\beta} = 1.838,\) \(\hat{\delta} = 0.170,\) \(\hat{\gamma} = 0.380\) | \(1.835,\) \(1.465,\) \(0.145,\) \(0.271\) |
| KMOFr  | \(\hat{\alpha} = 31946.7,\) \(\hat{\beta} = 0.607,\) \(\hat{\delta} = 4.806,\) \(\hat{\gamma} = 1.014,\) \(\hat{\lambda} = 13724.5\) | \(57.93,\) \(0.106,\) \(0.101,\) \(0.139,\) \(106.87\) |
| KFr    | \(\hat{\alpha} = 9378.5,\) \(\hat{\beta} = 0.084,\) \(\hat{\delta} = 5.513,\) \(\hat{\gamma} = 7160.5\) | \(80.43,\) \(0.024,\) \(17.49,\) \(36.07\) |
| EFr    | \(\hat{\alpha} = 1426.6,\) \(\hat{\beta} = 0.249,\) \(\hat{\delta} = 13.746\) | \(36.07,\) \(0.070,\) \(13.512\) |
| TEFr   | \(\hat{\alpha} = 1250.6,\) \(\hat{\beta} = 0.249,\) \(\hat{\delta} = 13.487,\) \(\hat{\gamma} = – 0.117\) | \(33.903,\) \(0.076,\) \(13.871,\) \(0.487\) |
| BExFr  | \(\hat{\alpha} = 0.116,\) \(\hat{\beta} = 4.364,\) \(\hat{\delta} = 0.043,\) \(\hat{\gamma} = 9.384,\) \(\hat{\lambda} = 6.405\) | \(0.026,\) \(0.025,\) \(0.048,\) \(2.709,\) \(0.969\) |
| Fr     | \(\hat{\alpha} = 7.865,\) \(\hat{\beta} = 0.694\) | \(2.091,\) \(0.091\) |

### Table 6: Goodness-of-fit statistics for dataset I

| Model  | \(-2L\) | AIC  | \(W^*\) | \(A^*\) | KS  | p-value |
|--------|---------|------|---------|---------|-----|---------|
| OChFr  | 300.183 | 308.183 | 0.0417 | 0.3148 | 0.105 | 0.863 |
| WFr    | 302.577 | 310.577 | 0.0613 | 0.4255 | 0.114 | 0.780 |
| KMOFr  | 304.804 | 314.804 | 0.0836 | 0.5469 | 0.140 | 0.533 |
| KFr    | 304.832 | 314.832 | 0.0946 | 0.6342 | 0.139 | 0.546 |
| EFr    | 307.788 | 313.788 | 0.1115 | 0.7050 | 0.135 | 0.581 |
| TEFr   | 307.760 | 315.760 | 0.1104 | 0.7006 | 0.136 | 0.569 |
| BExFr  | 309.905 | 319.905 | 0.1393 | 0.8549 | 0.140 | 0.536 |
| Fr     | 311.997 | 315.997 | 0.1601 | 0.9759 | 0.149 | 0.456 |

### Table 7: The MLEs with their SEs for dataset II
Table 8: Goodness-of-fit statistics for dataset II

| Model | \(-2L\)  | AIC | W   | A     | KS  | p value |
|-------|---------|-----|-----|-------|-----|---------|
| OChW  | 155.058 | 163.058 | 0.0111 | 0.0857 | 0.046 | 0.999   |
| OLLEW | 155.801 | 163.801 | 0.0159 | 0.1018 | 0.062 | 0.998   |
| ALTW  | 157.218 | 163.218 | 0.0154 | 0.1221 | 0.059 | 0.998   |
| BW    | 158.076 | 166.076 | 0.0210 | 0.1696 | 0.088 | 0.917   |
| MOW   | 162.640 | 168.64  | 0.0496 | 0.3766 | 0.092 | 0.889   |
| TCWG  | 162.640 | 170.640 | 0.0496 | 0.3766 | 0.092 | 0.889   |
| LIW   | 163.775 | 169.775 | 0.0636 | 0.4815 | 0.102 | 0.802   |
| W     | 164.951 | 168.951 | 0.0769 | 0.5730 | 0.108 | 0.742   |

Figure 5: The estimated CDFs for dataset I

Figure 6: The fitted PDFs for dataset I
Figure 7: The P-P plots for dataset I

Figure 8: The estimated CDFs for dataset II

Figure 9: The fitted PDFs for dataset II
Now, the different methods of estimation mentioned previously will be used to estimate the unknown parameters of the OChFr and OChW models. The KS statistic and its p value are provided to verify the best estimators. Tables 9 and 10 report the estimates of the unknown parameters using five estimation methods and the values of KS with corresponding p value for datasets I and II, respectively.

### Table 9: Various estimators of the OChFr parameters, KS and p-values for data set I

| Method | Estimator → | $\hat{\alpha}$ | $\hat{\beta}$ | $\hat{\xi}$ | $\hat{b}$ | KS  | p value |
|--------|-------------|----------------|---------------|-------------|---------|-----|---------|
| LSE    |             | 0.309          | 1.266         | 2.518       | 0.212   | 0.116 | 0.770   |
| WLSE   |             | 0.459          | 1.513         | 4.088       | 0.195   | 0.107 | 0.848   |
| MPSE   |             | 0.284          | 0.362         | 8.091       | 0.874   | 0.117 | 0.754   |
| CVME   |             | 0.157          | 0.873         | 1.259       | 0.288   | 0.113 | 0.791   |
| PCE    |             | 0.142          | 0.153         | 4.092       | 2.046   | 0.114 | 0.783   |

### Table 10: Various estimators of the OChW parameters, KS and p-values for dataset II

| Method | Estimator → | $\hat{\alpha}$ | $\hat{\beta}$ | $\hat{\xi}$ | $\hat{b}$ | KS  | p value |
|--------|-------------|----------------|---------------|-------------|---------|-----|---------|
| LSE    |             | 0.353          | 0.816         | 1.879       | 7.749   | 0.057 | 0.998   |
| WLSE   |             | 0.244          | 0.450         | 2.609       | 6.503   | 0.073 | 0.983   |
| MPSE   |             | 0.240          | 0.410         | 2.501       | 6.407   | 0.104 | 0.775   |
| CVME   |             | 0.007          | 0.118         | 0.544       | 0.059   | 0.057 | 0.998   |
| PCE    |             | 0.358          | 0.817         | 1.964       | 7.749   | 0.053 | 0.997   |

Figure 10: The P-P plots for dataset II
Tables 9 and 10 illustrate that all estimation methods work quite well. However, the MLE method gives the best estimation for the model parameters and consequently we recommend using it for estimating the model parameters for both datasets. Figures 11, 12 and 13 show the fitted PDFs, estimated CDFs and P-P plots for both datasets using the estimators in Tables 9 and 10.

Tables 11 and 12 list some numerical values of some reliability concepts for datasets I and II.

It is seen, from Table 11, that the RF and HRF decrease and the MRL increases with $t \to \infty$, whereas the RF and MRL decrease and the HRF increases with $t \to \infty$ as seen from Table 12.

**Figure 11:** The fitted densities (left panel) and estimated CDFs (right panel) of the OChFr distribution for various estimation methods for dataset I.

**Figure 12:** The fitted densities (left panel) and estimated CDFs (right panel) of the OChW distribution for various estimation methods for dataset II.
CONCLUSIONS

In this study, we proposed a new generator of distributions called the odd Chen-G (OCh-G) family. Several of its statistical properties have been derived. The special sub models of the OCh-G family are capable of modelling symmetric and positive as well as negative skewness datasets. Moreover, these sub models provide a wide variation in the shape of the hazard rate, including decreasing, increasing, unimodal and bathtub shapes, and consequently the generated model can be used in modelling various types of data. Two special cases of the OCh-G family, called the OCh-Frechet and OCh-Weibull models were studied. The model parameters are estimated using six different estimation methods, namely, the maximum likelihood estimators, least squares estimators, weighted least squares estimators, maximum product of spacings estimators, Cramér-von Mises, and others.
Mises estimators and percentile based estimators. The maximum likelihood estimation gives the best estimators for both OCh-Fréchet and OCh-Weibull models. Finally, the two special cases of the OCh-G class are applied to two real datasets from the medicine and engineering fields to illustrate the flexibility of the proposed family.

Acknowledgements

This publication was supported by the Deanship of Scientific Research at Prince Sattam bin Abdulaziz University, Alkharj, Saudi Arabia.

REFERENCES

Afify A.Z., Alizadeh M., Zayed M., Ramires T.G. & Louzada F. (2018). The odd log-logistic exponentiated Weibull distribution: regression modelling, properties and applications. Iranian Journal of Science and Technology: Transactions A: Science 42: 2273–2288. DOI: https://doi.org/10.1007/s40995-018-0524-x

Afify A.Z., Nofal Z.M. & Butt N.S. (2014). Transmuted complementary Weibull geometric distribution. Pakistan Journal of Statistics and Operation Research 10: 435–454. DOI: https://doi.org/10.1080/03610926.2018.90309

Alzaid M., Afify A. Z., Eliwa M. S., & Ali S. (2020). The odd log-logistic Lindley-G family of distributions: properties, Bayesian and non-Bayesian estimation with applications. Computational Statistics, 35(1): 281–308. DOI: https://doi.org/10.1007/s00180-019-00932-9

Alzaatreh A., Lee C. & Famoye F. (2013). A new method for generating families of continuous distributions. Metron 71: 63–79. DOI: https://doi.org/10.1007/s40300-013-0007-y

Amini M., Mir M.T.K. & Ahmadi J. (2014). Log-gamma-generated families of distributions. Statistics 48: 913–932. DOI: https://doi.org/10.1007/s10231-004

Bonferroni C. (1930). Elementi di statistica generale. Seebach, Firenze.

Chaubey Y.P. & Zhang R. (2015). An extension of Chen family of survival distributions with bathtub shape or increasing hazard rate function. Communications in Statistics: Theory and Methods 44: 4049–4064. DOI: http://doi.org/10.1080/03610926.2014.997357

Chen Z. (2000). A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. Statistics and Probability Letters 49: 155–161. DOI: https://doi.org/10.1016/S0167-7152(00)00044-4

Cordeiro G.M., Afify A.Z., Ortega E.M., Suzuki A.K. & Mead M.E. (2019). The odd Lognormal generator of distributions: properties, estimation and applications. Journal of Computational and Applied Mathematics 347: 222–237. DOI: https://doi.org/10.1016/j.cam.2018.08.008

Cordeiro G.M., Afify A.Z., Yusof H.M., Cakmakyan S. & Olz G. (2018). The Lindley Weibull distribution: properties and applications. Anais da Academia Brasileira De Ciências 90: 2579–2589. DOI: http://dx.doi.org/10.1590/0001-3765201820170635

Dey S., Kumar D., Ramos P.L. & Louzada F. (2017). Exponentiated Chen distribution: properties and estimation. Communications in Statistics - Simulation and Computation 46(10): 8118–8139. DOI: https://doi.org/10.1080/03610918.2016.1267752

Elbatal I., Asha G. & Raja V. (2014). Transmuted exponentiated Fréchet distribution: properties and applications. Journal of Statistics Applications and Probability 3: 379–394. DOI: https://doi.org/10.12785/jasp/030309

El-Bassiouny A.H., EL-Damcese M., Abdelfattah M. & Eliwa M.S. (2017). Exponentiated generalized Weibull-Gompertz distribution with application in survival analysis. Journal of Statistics Applications and Probability 6: 7–16. DOI: https://doi.org/10.18576/jasp/060102

El-Bassiouny A.H., EL-Damcese M., Abdelfattah M. & Eliwa M.S. (2016). Mixture of exponentiated generalized Weibull-Gompertz distribution and its applications in reliability. Journal of Statistics Applications and Probability 5: 455–468. DOI: https://doi.org/10.18576/jasp/050310

El-Gohary A., EL-Bassiouny A.H. & EL-Morshed A.M. (2015a). Exponentiated flexible Weibull extension distribution. International Journal of Mathematics and its Applications 3(3-A): 1–12.

El-Gohary A., EL-Bassiouny A.H. & EL-Morshed A.M. (2015b). Inverse flexible Weibull extension distribution. International Journal of Computer Applications 115: 46–51. DOI: https://doi.org/10.5120/20127-2211

El-Morsy M., EL-Bassiouny A.H. & EL-Gohary A. (2017). Exponentiated inverse flexible Weibull extension distribution. Journal of Statistics Applications and Probability 6: 169–183. DOI: https://doi.org/10.18576/jasp/060114

El-Morsy M. & Eliwa M.S. (2019). The odd flexible Weibull-H family of distributions: properties and estimation with applications to complete and upper record data. Filomat 33: 2635–2652. DOI: https://doi.org/10.2298/FIL1909635E

Eliwa M.S., EL-Morsy M. & Ibrahim M. (2018). Inverse Gompertz distribution: properties and different estimation methods with application to complete andensored data. Annals of Data Science 6(2): 321–339. DOI: https://doi.org/10.1007/s40745-018-0173-0

Feigl P. & Zelen M. (1965). Estimation of exponential probabilities with concomitant information. Biometrics 21: 826–838.

Fréchet M. (1924). Sur la loi des erreurs d’observation. Bulletin de la Societ et e Mathematique de Moscou 33: 5–8.
Ghitany M.E., Al-Hussaini E.K. & Al-Jarallah R.A. (2005). Marshall-Olkin extended Weibull distribution and its application to censored data. *Journal of Applied Statistics* **32**: 1025–1034.
DOI: https://doi.org/10.1080/02664760500165008

Haghbin H., Ozel G., Alizadeh M. & Hamedani G.G. (2016). A new generalized odd log-logistic family of distributions. *Communications in Statistics: Theory and Methods* **46**: 9897–9920.
DOI: https://doi.org/10.1080/03610926.2016.1222428

Jehhan A., Mohamed I., Eliwa M.S., Al-mualim S. & Yousof H.M. (2018). The two-parameter odd Lindley Weibull lifetime model with properties and applications. *International Journal of Statistics and Probability* **7**(4): 57–68.
DOI: https://doi.org/10.5539/ijsp.v7n4p57

Lee C., Famoye F. & Olumolade O. (2007). Beta-Weibull distribution: some properties and applications to censored data. *Journal of Modern Applied Statistical Methods* **6**: 17–30.

Mead M.E., Afify A.Z., Hamedan G.G. & Ghosh I. (2017). The beta exponential Fréchet distribution with applications. *Australian Journal of Statistics* **46**: 41–63.
DOI: https://doi.org/10.17713/ajs.v46i.144

Mead M.E. & Abd-Eltawab A.R. (2014). A note on Kumaraswamy Fréchet distribution. *Australian Journal of Basic and Applied Sciences* **8**: 294–300.

Nadarajah S. & Kotz S. (2003). The exponentiated Fréchet distribution. *Interstat Electronic Journal* **20**: 1–7.

Nassar M., Afify A.Z., Dey S. & Kumar D. (2018). A new extension of Weibull distribution: properties and different methods of estimation. *Journal of Computational and Applied Mathematics* **336**: 439–457.
DOI: https://doi.org/10.1016/j.cam.2017.12.001

Shuaib M., Robert K. & Irene L. (2018). Kumaraswamy exponentiated Chen distribution for modelling lifetime data. *Applied Mathematics and Information Sciences* **12**: 617–623.
DOI: https://doi.org/10.18576/amis/paper*25-4-2018

Weibull W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics: Transactions, ASME* **18**: 293-297.

Xu K., Xie M., Tang L.C. & Ho S.L. (2003). Application of neural networks in forecasting engine systems reliability. *Applied Soft Computing* **2**: 255-268.
DOI: https://doi.org/10.1016/S1568-4946(02)00059-5