Journal of Statistics and Computer Science
Vol. 2, Issue 1, 2023, pp. 1-16
https://doi.org/10.47509/JSCS.2023.v02i01.01

Estimation of Matusita Overlapping Coefficient ρ for Two Weibull Distributions

Omar M. Eidous and **Hala K. Maqableh**

omarm@yu.edu.jo halakhaled1711@gmail.com

Department of Statistics - Faculty of Science - Yarmouk University, Irbid - Jordan

Abstract

The Matusita overlapping coefficient (MOC) ρ is defined as the similarity or agreement between two distributions. Let $f_X(x)$ and $f_Y(y)$ be two probability density functions for the two independent continuous random variables X and Y respectively, then the MOC is defined by $\rho = \int \sqrt{f_X(x) f_Y(x)}\, dx$. Some studies estimated ρ under pair Weibull distributions with different scale parameters and the same shape parameter. Without using this assumption, it is difficult to find the mathematical formula of ρ. This paper deals with the estimation of ρ under pair Weibull distributions without any restrictions on the parameters of the Weibull distributions. A new technique is suggested to estimate ρ, which can be used with and without using any assumptions about these parameters. In all situations, the maximum likelihood method is used to estimate the Weibull distributions parameters. The properties of the resulting proposed new estimators of ρ are investigated and compared with some existing parametric and nonparametric kernel estimators via Monte-Carlo simulation technique. The results show that the new technique is very competitor and the performances of the resulting new estimators are better than that of the nonparametric kernel estimators for all considered cases.

Keywords: Matusita Overlapping Coefficient; Maximum Likelihood Method; Weibull Distribution; Relative Bias and Relative Root Mean Square Error.

MSC: 62G10

Article History
Received : 16 February 2023; Revised : 29 April 2023, June 13, 2023; Accepted : 13 May 2023; Published : 30 June 2023

To cite this paper
Omar M. Eidous & Hala K. Maqableh (2023). Estimation of Matusita Overlapping Coefficient ρ for Two Weibull Distributions. *Journal of Statistics and Computer Science.* 2(1), 1-16.

1. Introduction

To determine the similarity between two populations in statistics, we can use the Matusita overlapping coefficient (MOC) ρ. MOC is defined as a measure of agreement between two probability distributions. Let $f_X(x)$ and $f_Y(y)$ be two probability density functions for the two independent continuous random variables X and Y respectively, the MOC is defined by (Matusita, 1955),

$$\rho = \int \sqrt{f_X(x) f_Y(x)}\, dx$$

If the value of ρ is 1 (i.e. $\rho = 1$) then $f_X(x) = f_Y(x)$. If $\rho = 0$ then the supports of the two densities $f_X(x)$ and $f_Y(x)$ have no interior points in common.

In general, there are three main well-known measures of overlapping, namely: Matusita measure ρ, Morisita measure λ and Weitzman measure Δ. However, we have paid attention to the Matusita measure ρ in this article. Overlapping measures are applied in different areas like, genetic (Federer, 1963), ecology (Pianka, 1973), income (Gastwirth, 1975), reliability analysis to estimate the proportion of machines or electronic devices that have similar range of failure time (Ichikawa, 1993 and Dhaker et al., 2019). Dhaker et al. (2019) reported: "The machines may come from two different sources or may be under different stress, which implies different probability densities of failure time. This proportion can be measured by the MOC coefficient of the two densities".

In the literature, the main two methods to estimate the overlapping measures are the parametric and the nonparametric methods. On one hand, the nonparametric method can be used when the shape of the densities $f_1(x)$ and $f_2(x)$ for specific data sets is difficult to determine. In this case, $f_1(x)$ and $f_2(x)$ are estimated by using the non-parametric methods such as Jackknife and Bootstrap (Mulekar and Mishra, 1994), kernel method (Eidous and Al-Talafha. 2022 and Alodat et *al.*, 2022) and empirical distribution method (Pastore, 2018). Despite of recent advances in nonparametric methods, parametric methods are still widely used among many researchers (other than statisticians) in different areas, mainly because of their simplicity. In addition, the parametric method is better than the corresponding nonparametric counterparts when all required assumptions are met. Parametric methods assume that the shapes of probability density functions $f_1(x)$ and $f_2(x)$ are known, and they depend on an unknown parameter θ, which may be a vector. The parameter(s) θ can be estimated by the usual point estimation methods such as the maximum likelihood (ML) method or the method of moments (MM).

Inman and Bradly (1989) estimated overlapping measures under the normality assumption of both distributions with equal variances. In addition, they investigated some cases where the overlapping measures can be applied. Mulekar and Mishra (1994) derived and studied the overlapping measures of two normal distributions with equal means and different variances. Independently, Eidous and Al-Daradkeh (2022) and Eidous and Shourman (2022) each presented novel techniques to estimate the Matusita overlapping coefficient ρ under pair normal distributions without using any restrictions on the equality of their location parameters and the equality of their scale parameters. Al-Saidy et *al.* (2005) derived the formulas of overlapping coefficients in the case of Weibull distributions and estimate them under the conditions that the two shape parameters are equal. The pdf of Weibull model with a scale parameter α and a shape parameter β is,

$$f_X(x) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-(x/\alpha)^{\beta}}, \qquad x > 0, \qquad \alpha, \beta > 0.$$

We will denote this by $X \sim W(x; \alpha,\ \beta)$. Figure (1) displays the plots of Weibull distribution with shape parameter $\beta = 2$ and four values for the scale parameter $\alpha = 4, 5, 6, 8,$ while Figure (2) plots four Weibull distribution curves with $\alpha = 1$ and $\beta = 2, 3, 4, 6$. The Weibull distribution is associated with a number of density functions. In particular, it interpolates between the exponential distribution ($\beta = 1$) and the Rayliegh distribution ($\beta = 2$ and $\alpha = \sqrt{2}\sigma$). Let $X \sim f_X(x) = W_1(x; \alpha_1, \beta_1)$ and $Y \sim f_Y(y) = W_2(y; \alpha_2, \beta_2)$ where X and Y are independent random variables. Assume that $\beta_1 = \beta_2 = (\beta$ say), $K = \alpha_1/\alpha_2$ and $Q = (2\beta - 1)/\beta$, the formulas of the measures ρ based on the Weibull models that derived by Al-Saidy et al., (2005) is,

$$\rho = \int_0^{\infty} \sqrt{W_1(x; \alpha_1, \beta\,) W_2(x; \alpha_2, \beta\,)}dx = \frac{2\sqrt{K^{\beta}}}{1 + K^{\beta}}.$$

Let $\hat{K}$ and $\hat{\beta}$ be the maximum likelihood (ML) estimator of K and β respectively, then

$$\hat{\rho}_{Al} = \frac{2\sqrt{\hat{K}^{\hat{\beta}}}}{1 + \hat{K}^{\hat{\beta}}}$$

is the ML estimator of ρ. The value of ρ for two exponential distributions can be obtained by taking $\beta = 1$ in the Weibull distribution. The case of two exponential distributions was studied by [15, 16]. Samawi and al-Saleh (2008) studied the effect of sampling scheme on overlapping coefficients. To estimate ρ, Al-Saidy *et al*. (2005) used the maximum likelihood method to estimate the corresponding parameters α_1, α_2 and β.

If $\beta_1 = \beta_2 = \beta$ then the closed form of the interested parameter ρ can be obtained, which is as given by the above formula. We can estimate it by estimating the corresponding parameters α_1, α_2 and β by using − for example − maximum likelihood method. Now if $\alpha_1 = \alpha_2$ but $\beta_1 \neq \beta_2$ then the closed form of ρ cannot be obtained. The same thing can be said if $\alpha_1 \neq \alpha_2$ and $\beta_1 \neq \beta_2$. Therefore, for the last two cases, we cannot estimate ρ by the same technique as of the first case. The main aim of this article is to suggest a new technique that enables us to estimate ρ for all of the above different cases including the first one.

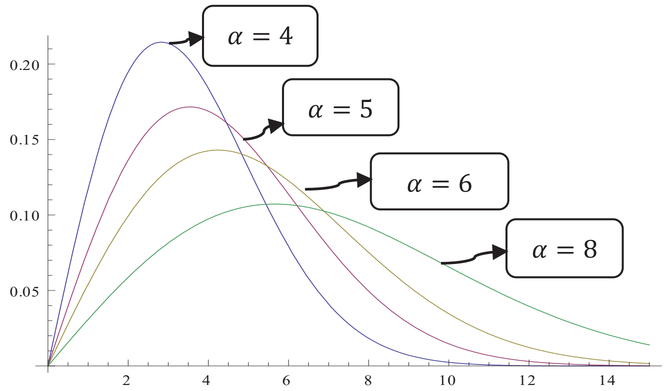

Figure 1. The plot of $W(x;\alpha,2)$, $\alpha = 4, 5, 6, 8$.

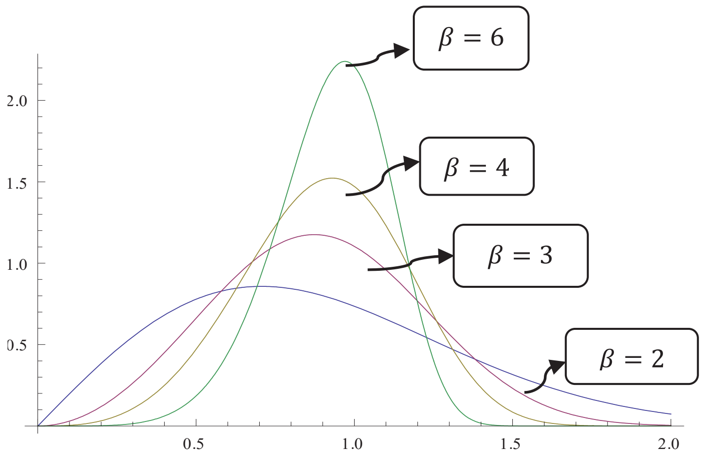

Figure 2. The plot of $W(x;1,\beta)$, $\beta = 2, 3, 4, 6$.

2. Estimation of Weibull Distributions Parameters

To estimate the Matusita Measure ρ we need to estimate the parameters of the two Weibull distributions. We suggest using the well-known maximum likelihood method, which produces consistent estimators for the various parameters.

Let $X_1, X_2, \dots, X_{n_1}$ be a random sample of size n_1 from $W_1(x; \alpha_1, \beta_1)$ and let $Y_1, Y_2, \dots, Y_{n_2}$ be another random sample of size n_2 from $W_2(y; \alpha_2, \beta_2)$, where the two samples are independent. The likelihood function of $X_1, X_2, \dots, X_{n_1}$ and $Y_1, Y_2, \dots, Y_{n_2}$ is,

$$L(\alpha_1, \alpha_2, \beta_1, \beta_2) = \frac{{\beta_1}^{n_1}}{{\alpha_1}^{n_1\beta_1}} \frac{{\beta_2}^{n_2}}{{\alpha_2}^{n_2\beta_2}} \left(\prod_{i=1}^{n_1} x_i\right)^{\beta_1-1} \left(\prod_{i=1}^{n_2} y_i\right)^{\beta_2-1} e^{-\frac{1}{\alpha_1^{\beta_1}}\sum_{i=1}^{n_1} x_i^{\beta_1} - \frac{1}{\alpha_2^{\beta_2}}\sum_{i=1}^{n_2} y_i^{\beta_2}}$$

In the following subsections, the maximum likelihood estimators (MLEs) of the different Weibull distribution parameters are obtained by considering the three cases:

a) The two scale parameters are equal

b) The two shape parameters are equal and

c) There is no restrictions about the scale and shape parameters.

2.1 Two Scale Parameters are Equal

Suppose that the two scale parameters are equal, i.e. $\alpha_1 = \alpha_2 = (\alpha \text{ say})$, then the likelihood function becomes,

$$L(\alpha, \beta_1, \beta_2) = \frac{{\beta_1}^{n_1}{\beta_2}^{n_2}}{\alpha^{n_1\beta_1 + n_2\beta_2}} \left(\prod_{i=1}^{n_1} x_i\right)^{\beta_1-1} \left(\prod_{i=1}^{n_2} y_i\right)^{\beta_2-1} e^{-\frac{1}{\alpha^{\beta_1}}\sum_{i=1}^{n_1} x_i^{\beta_1} - \frac{1}{\alpha^{\beta_2}}\sum_{i=1}^{n_2} y_i^{\beta_2}}$$

The ML estimators of α, β_1 and β_2 can be obtained by solving the following three equations numerically (such as Newton-Raphson method) with respect to α, β_1 and β_2,

$$\frac{\partial lnL(\alpha, \beta_1, \beta_2)}{\partial \alpha} = -\frac{n_1\beta_1 + n_2\beta_2}{\alpha} + \beta_1 \frac{\sum_{i=1}^{n_1} {x_i}^{\beta}}{\alpha^{\beta_1+1}} + \beta_2 \frac{\sum_{i=1}^{n_2} {y_i}^{\beta_2}}{\alpha^{\beta_2+1}} = 0$$

$$\frac{\partial lnL(\alpha, \beta_1, \beta_2)}{\partial \beta_1} = \frac{n_1}{\beta_1} - n_1 \, ln\alpha + \sum_{i=1}^{n_1} ln \, x_i - \sum_{i=1}^{n_1} \left[\left(\frac{x_i}{\alpha}\right)^{\beta_1} ln\left(\frac{x_i}{\alpha}\right)\right] = 0$$

$$\frac{\partial lnL(\alpha, \beta_1, \beta_2)}{\partial \beta_2} = \frac{n_2}{\beta_2} - n_2 \, ln\alpha + \sum_{i=1}^{n_2} ln \, y_i - \sum_{i=1}^{n_2} \left[\left(\frac{y_i}{\alpha}\right)^{\beta_2} ln\left(\frac{y_i}{\alpha}\right)\right] = 0.$$

2.2 Two Shape Parameters are Equal

Suppose that the two shape parameters are equal, i.e. $\beta_1 = \beta_2 = (\beta$ say), then the likelihood function becomes,

$$L(\alpha_1, \alpha_2, \beta) = \frac{\beta^{n_1+n_2}}{{\alpha_1}^{n_1\beta}{\alpha_2}^{n_2\beta}} \left(\prod_{\mathrm{i}=1}^{n_1} x_i \prod_{\mathrm{i}=1}^{n_2} y_i \right)^{\beta-1} e^{-\frac{1}{\alpha_1{}^\beta}\sum_{i=1}^{n_1} \mathrm{x}_\mathrm{i}^\beta - \frac{1}{\alpha_2{}^\beta}\sum_{i=1}^{n_2} \mathrm{y}_\mathrm{i}^\beta},$$

The ML estimators of α_1, α_2 and β are obtained by solving the following three equations numerically using Newton-Raphson method,

$$\frac{\partial lnL(\alpha_1, \alpha_2, \beta)}{\partial \alpha_1} = \frac{-n_1\beta}{\alpha_1} + \frac{\beta \sum_{i=1}^{n_1} x_i^\beta}{{\alpha_1}^{\beta+1}} = 0$$

$$\frac{\partial lnL(\alpha_1, \alpha_2, \beta)}{\partial \alpha_2} = \frac{-n_2\beta}{\alpha_2} + \frac{\beta \sum_{i=1}^{n_2} y_i^\beta}{{\alpha_2}^{\beta+1}} = 0$$

$$\frac{\partial lnL(\alpha_1, \alpha_2, \beta)}{\partial \beta} = \frac{n_1 + n_2}{\beta} - \ (n_1 ln\alpha_1 + n_2 ln\alpha_2) + \left(\sum_{i=1}^{n_1} lnx_i + \sum_{i=1}^{n_2} lny_i \right)$$
$$- \sum_{i=1}^{n_1} \left[\left(\frac{x_i}{\alpha_1} \right)^\beta ln \left(\frac{x_i}{\alpha_1} \right) \right] - \sum_{i=1}^{n_2} \left[\left(\frac{y_i}{\alpha_2} \right)^\beta ln \left(\frac{y_i}{\alpha_2} \right) \right] = 0.$$

2.3 Scale and Shape Parameters are not Necessarily Equal

Without using any restrictions about the Weibull distributions parameters, the likelihood function $L(\alpha_1, \alpha_2, \beta_1, \beta_2)$ is as given above. Therefore, the ML estimators of $\alpha_1, \alpha_2, \beta_1$ and β_2 are obtained by solving the following equations simultaneously with respect to $\alpha_1, \alpha_2, \beta_1$ and β_2.

$$\frac{\partial lnL(\alpha_1, \alpha_2, \beta_1, \beta_2)}{\partial \alpha_1} = \frac{-n_1\beta_1}{\alpha_1} + \frac{\beta_1 \sum_{i=1}^{n_1} x_i^{\beta_1}}{{\alpha_1}^{\beta_1+1}} = 0$$

$$\frac{\partial lnL(\alpha_1, \alpha_2, \beta_1, \beta_2)}{\partial \alpha_2} = \frac{-n_2\beta_2}{\alpha_2} + \frac{\beta_2 \sum_{i=1}^{n_2} y_i^{\beta_2}}{{\alpha_2}^{\beta_2+1}} = 0$$

$$\frac{\partial lnL(\alpha_1, \alpha_2, \beta_1, \beta_2)}{\partial \beta_1} = \frac{n_1}{\beta_1} - n_1\, ln\alpha_1 + \sum_{i=1}^{n_1} ln\, x_i - \sum_{i=1}^{n_1} \left[\left(\frac{x_i}{\alpha_1} \right)^{\beta_1} ln \left(\frac{x_i}{\alpha_1} \right) \right] = 0$$

$$\frac{\partial lnL(\alpha_1, \alpha_2, \beta_1, \beta_2)}{\partial \beta_2} = \frac{n_2}{\beta_2} - n_2\, ln\alpha_2 + \sum_{i=1}^{n_2} ln\, y_i - \sum_{i=1}^{n_2} \left[\left(\frac{y_i}{\alpha_2} \right)^{\beta_2} ln \left(\frac{y_i}{\alpha_2} \right) \right] = 0.$$

Again Newton-Raphson method can be used to solve the above equations.

3. Estimation of Matusita Measure ρ

The formula of Matusita Measure ρ under two independent Weibull distributions is defined as,

$$\rho = \int_0^\infty \sqrt{f_X(x) f_Y(x)}\, dx = \int_0^\infty \sqrt{W_1(x;\alpha_1,\beta_1) W_2(x;\alpha_2,\beta_2)} dx$$

To be able to estimate ρ we will write it as an expected value of some functions. By considering $W_2(X;\alpha_2,\beta_2)/W_1(X;\alpha_1,\beta_1)$ as a function of X and $W_1(Y;\alpha_1,\beta_1)/W_2(Y;\alpha_2,\beta_2)$ as a function of Y, then,

$$\begin{aligned} E\left(\frac{W_2(X;\alpha_2,\beta_2)}{W_1(X;\alpha_1,\beta_1)}\right)^{\frac{1}{2}} &= \int_0^\infty \left(\frac{W_2(x;\alpha_2,\beta_2)}{W_1(x;\alpha_1,\beta_1)}\right)^{\frac{1}{2}} W_1(x;\alpha_1,\beta_1)\, dx \\ &= \int_0^\infty \sqrt{W_1(x;\alpha_1,\beta_1) W_2(x;\alpha_2,\beta_2)} dx \\ &= \rho, \end{aligned}$$

and

$$\begin{aligned} E\left(\frac{W_1(Y;\alpha_1,\beta_1)}{W_2(Y;\alpha_2,\beta_2)}\right)^{\frac{1}{2}} &= \int_0^\infty \left(\frac{W_1(y;\alpha_1,\beta_1)}{W_2(y;\alpha_2,\beta_2)}\right)^{\frac{1}{2}} W_2(y;\alpha_2,\beta_2)\, dy \\ &= \int_0^\infty \sqrt{W_1(y;\alpha_1,\beta_1) W_2(y;\alpha_2,\beta_2)} dy \\ &= \int_0^\infty \sqrt{W_1(x;\alpha_1,\beta_1) W_2(x;\alpha_2,\beta_2)} dx \\ &= \rho. \end{aligned}$$

Also, ρ can be expressed as the average of the above two quantities as follows,

$$\rho = \frac{1}{2}\left[E\left(\frac{W_2(X;\alpha_2,\beta_2)}{W_1(X;\alpha_1,\beta_1)}\right)^{\frac{1}{2}} + E\left(\frac{W_1(Y;\alpha_1,\beta_1)}{W_2(Y;\alpha_2,\beta_2)}\right)^{\frac{1}{2}}\right].$$

Therefore, ρ can be expressed as follows,

$$\rho = E\left(\frac{W_2(X;\alpha_2,\beta_2)}{W_1(X;\alpha_1,\beta_1)}\right)^{\frac{1}{2}} = E\left(\frac{W_1(Y;\alpha_1,\beta_1)}{W_2(Y;\alpha_2,\beta_2)}\right)^{\frac{1}{2}} = \frac{1}{2}\left[E\left(\frac{W_2(X;\alpha_2,\beta_2)}{W_1(X;\alpha_1,\beta_1)}\right)^{\frac{1}{2}} + E\left(\frac{W_1(Y;\alpha_1,\beta_1)}{W_2(Y;\alpha_2,\beta_2)}\right)^{\frac{1}{2}}\right].$$

Let $X_1, X_2, \ldots, X_{n_1}$ be a random sample of size n_1 from $W_1(x;\alpha_1,\beta_1)$ and let $Y_1, Y_2, \ldots, Y_{n_2}$ be another random sample of size n_2 from $W_2(y;\alpha_2,\beta_2)$, where the two samples are independent. We suggest to estimate $E\left(\frac{W_2(X;\alpha_2,\beta_2)}{W_1(X;\alpha_1,\beta_1)}\right)^{\frac{1}{2}}$ by the mean of

$\left(\frac{W_2(X_i;\alpha_2,\beta_2)}{W_1(X_i;\alpha_1,\beta_1)}\right)^{\frac{1}{2}}$, $i = 1,2,\dots,n_1$ and to estimate $E\left(\frac{W_1(Y;\alpha_1,\beta_1)}{W_2(Y;\alpha_2,\beta_2)}\right)^{\frac{1}{2}}$ by the mean of $\left(\frac{W_1(Y_j;\alpha_1,\beta_1)}{W_2(Y_j;\alpha_2,\beta_2)}\right)^{\frac{1}{2}}$, $j = 1,2,\dots,n_2$. In the same way, we can estimate $\frac{1}{2}\left[E\left(\frac{W_2(X;\alpha_2,\beta_2)}{W_1(X;\alpha_1,\beta_1)}\right)^{\frac{1}{2}} + E\left(\frac{W_1(Y;\alpha_1,\beta_1)}{W_2(Y;\alpha_2,\beta_2)}\right)^{\frac{1}{2}}\right]$ by the average of the above two means. However, the resulting estimators by adopting this technique are still depend on unknown quantities, which are the parameters α_1, α_2, β_1 and β_2. Suppose that the maximum likelihood estimators of α_1, α_2, β_1 and β_2 are $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\beta}_1$ and $\hat{\beta}_2$ respectively (see Section 2), then the suggested three estimators of ρ are,

$$\hat{\rho}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1}\left(\frac{W_2(X_i;\hat{\alpha}_2,\hat{\beta}_2)}{W_1(X_i;\hat{\alpha}_1,\hat{\beta}_1)}\right)^{\frac{1}{2}}$$

$$\hat{\rho}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2}\left(\frac{W_1(Y_j;\hat{\alpha}_1,\hat{\beta}_1)}{W_2(Y_j;\hat{\alpha}_2,\hat{\beta}_2)}\right)^{\frac{1}{2}}$$

and

$$\hat{\rho}_3 = \frac{1}{2}\left[\frac{1}{n_1}\sum_{j=1}^{n_1}\left(\frac{W_2(X_i;\hat{\alpha}_2,\hat{\beta}_2)}{W_1(X_i;\hat{\alpha}_1,\hat{\beta}_1)}\right)^{\frac{1}{2}} + \frac{1}{n_2}\sum_{j=1}^{n_2}\left(\frac{W_1(Y_j;\hat{\alpha}_1,\hat{\beta}_1)}{W_2(Y_j;\hat{\alpha}_2,\hat{\beta}_2)}\right)^{\frac{1}{2}}\right].$$

As a preliminary simulation study shown, the last estimator $\hat{\rho}_3$ of ρ is more stable than the other two estimators $\hat{\rho}_1$ and $\hat{\rho}_2$. Therefore, only the performance of $\hat{\rho}_3$ is investigated in our simulation study in Section (4).

4. Simulation Study

To study the performance of the proposed estimator, a simulation study is performed. To simulate the data and to be consistent with the derivations of this paper, we considered the following three cases:

Case (1). A pair of Weibull distribution have the same scale parameters (i.e. $\alpha_1 = \alpha_2 = \alpha$).

Case (2). A pair of Weibull distribution have the same shape parameters (i.e. $\beta_1 = \beta_2 = \beta$)

Case (3). A pair of Weibull distributions have different scale parameters and different shape parameters (i.e. $\alpha_1 \neq \alpha_2$ and $\beta_1 \neq \beta_2$).

We simulate $x_1, x_2, \dots, x_{n_1}$ from $f_X(x) = W_1(x;\alpha_1,\beta_1)$ and $y_1, y, \dots, y_{n_2}$ from $f_Y(y) = W_2(y;\alpha_2,\beta_2)$ with specific values of the scale and shape parameters. To cover the most possible cases in practice, 4 pair of distributions are chosen for each case. The selected parameters $\alpha_1, \alpha_2, \beta_1, \beta_2$ and the exact value of ρ for each selection are given in Table (1). The samples size are chosen to be $(n_1, n_2) = (10{,}10), (20{,}30), (50{,}50), (100{,}200)$.

In the first case (Case 1), we studied two proposed estimators for ρ. The first proposed estimator is the estimator that took into account the equality of the scale parameters (see Subsection 3.1), we will denote it by $\hat{\rho}_{P1}$ and the second estimator is $\hat{\rho}_3$, which was developed without any restrictions on the distribution parameters (see Subsection 3.3). Similarly, we considered two proposed estimators in the second case (Case 2). The first estimator will be denoted by $\hat{\rho}_{P2}$ (take into account the equality of the shape parameters, see Subsection 2.2) and the second estimator is $\hat{\rho}_3$. In the third case (Case `3), the only proposed estimator is $\hat{\rho}_3$.

In the case of equal shape parameters of Weibull distributions (Case 2), the estimator $\hat{\rho}_{Al}$ is also considered for sake of comparison (see Section 1). In addition, the nonparametric kernel estimator $\hat{\rho}_k$ (see Eidous and Al-Talafha, 2022) is also included in this study. The kernel estimator $\hat{\rho}_k$ can be used for the above three cases since it requires no assumptions about the distribution itself (Eidous and Al-Talafha, 2022).

For each estimator, we compute the Relative Bias (RB) and Relative Root Mean Square Error (RRMSE) based on $R = 1000$ replication. RB and RRMSE are defined as follows,

$$RB = \frac{\hat{E}(estimator) - exact\ \rho}{exact\ \rho},$$

and

$$RRMSE = \frac{\sqrt{\widehat{MSE}(estimator)}}{exact\ \rho}$$

For example, if $\hat{\rho}$ is the estimator of ρ and if $\hat{\rho}_{(j)}$ is the value of $\hat{\rho}$ computed based on a sample iteration $j, j = 1, 2, \dots, R = 1000$ then,

$$\hat{E}(\hat{\rho}) = \sum_{j=1}^{1000} \hat{\rho}_{(j)} / 1000$$

and

$$\widehat{MSE}(\hat{\rho}) = \sum_{j=1}^{1000} \left(\hat{\rho}_{(j)} - \hat{E}(\hat{\rho})\right)^2 / 1000$$

5. Simulation Results

Tables (2), (3) and (4) contain values of RB and RRMSE for the different estimators. The results of Table (2) depend on the data simulated from pair Weibull distributions with the equal scale parameters. Table (3) contains the results when the data are simulated from pair Weibull distribution with equal shape parameters, while the results of Table (4) depends on the data simulated from pair Weibull distributions with different scale and different shape parameters. The two estimators $\hat{\boldsymbol{\rho}}_{\boldsymbol{k}}$ (kernel estimator) and $\hat{\boldsymbol{\rho}}_{\boldsymbol{3}}$ (proposed estimator without any restrictions on the Weibull distribution parameters) have been derived without any restrictions on the Weibull parameters. Therefore, the results of these two estimators are included in the three tables. The results related the proposed estimator $\hat{\boldsymbol{\rho}}_{\boldsymbol{P1}}$ are presented only in Table (2) since this estimators was derived under the assumption that the two scale parameters are equal. The two estimators $\hat{\boldsymbol{\rho}}_{\boldsymbol{Al}}$ and $\hat{\boldsymbol{\rho}}_{\boldsymbol{P2}}$ (proposed estimator) were derived when the two shape parameters are assumed to be equal. Therefore, their results are included only in Table (3).

From these simulation results, we can conclude the following:

1. It is obvious that $|RB|$s associated with the kernel estimator $\hat{\rho}_k$ are significantly large compared with the other estimators, especially for small samples sizes. All values RBs values of the kernel estimates are negative, which indicates that -on the average- $\hat{\rho}_k$ is underestimate the true value ρ.
2. As the sample sizes increases the $RRMSE$ of the different estimators decreases. This is a good sign for the consistency of the different estimators that considered in this study.
3. The values of $RRMSE$ for the proposed estimators in different cases indicate that their performances are much better than the kernel estimator in all considered cases.
4. As the results of Table (3) shown, the estimator $\hat{\rho}_{Al}$ performs better than the suggested estimators for small sample sizes and small exact value of ρ. This result can be seen by examining the corresponding values of $RRMSE$ despite that the $|RB|$s associated with the proposed estimators are less than that associated with $\hat{\rho}_{Al}$ in many cases . If the sample sizes gets large (say, $n_1, n_1 \geq 50$) the performances of $\hat{\rho}_{Al}$ and the proposed estimators are similar in most considered cases. The major disadvantage of $\hat{\rho}_{Al}$ is that it can be used only when the data are assumed to follow pair Weibull distributions with equal shape parameters, which is often not feasible in practice. Of course, this limitation reduces the importance of this estimator as a general estimator for ρ.

5. Based on the values of $RRMSE$s in Table (2) and by comparing the performances of the two proposed estimators $\hat{\rho}_{P1}$ and $\hat{\rho}_3$ together; we can see that their performances seem to be similar (especially for large sample sizes) in most considered cases despite that $\hat{\rho}_{P1}$ is a bit better than $\hat{\rho}_3$. Similarly, we can see that $\hat{\rho}_{P2}$ is a bit better than $\hat{\rho}_3$ based on the results of Table (3).

6. Despite the previous conclusion and by taking into account that $\hat{\rho}_3$ is developed without any constraints on the Weibull distribution parameters, then we can recommend $\hat{\rho}_3$ as a general estimator for ρ in the case of Weibull distributions.

Table 1. Exact values of the overlapping coefficient ρ for the 12 simulated pairs of Weibull distributions

Weibull distributions		$f_X(x)$	$f_Y(y)$	ρ
Case 1: Equal scales	A	$W(x;1,3)$	$W(y;1,4)$	0.9816
	B	$W(x;1,3)$	$W(y;1,6.2)$	0.8971
	C	$W(x;1,3)$	$W(y;1,10.3)$	0.7581
	D	$W(x;1,3)$	$W(y;1,20.4)$	0.5679
Case 2: Equal shapes	A	$W(x;1,3)$	$W(y;1.2,3)$	0.9637
	B	$W(x;1,3)$	$W(y;1.5,3)$	0.8398
	C	$W(x;1,3)$	$W(y;1.8,3)$	0.7069
	D	$W(x;1,3)$	$W(y;3.5,3)$	0.2984
Case 3: Different scales and different shapes	A	$W(x;1,1.2)$	$W(y;2,1.8)$	0.9793
	B	$W(x;1,1.5)$	$W(y;3,1.9)$	0.8435
	C	$W(x;1,1.8)$	$W(y;4,2.1)$	0.6893
	D	$W(x;1,6)$	$W(y;3,2)$	0.3929

Table 2. The RB and RRMSE of the three estimators $\hat{\rho}_k$, $\hat{\rho}_{p_1}$ and $\hat{\rho}_{p_2}$ when the data are simulated from pair Weibull distributions with equal scale parameters ($\alpha_1 = \alpha_2 = 1$).

$\rho_{exact} = 0.9816$

(β_1, β_2)	(n_1, n_2)		$\hat{\rho}_k$	$\hat{\rho}_3$	$\hat{\rho}_{P1}$
(3, 4)	(10, 10)	RB	-0.3490	-0.0635	-0.0223
		RRMSE	0.3692	0.0894	0.0595
	(20, 30)	RB	-0.1786	-0.0262	-0.0085
		RRMSE	0.1956	0.0531	0.0470
	(50, 50)	RB	-0.1204	-0.0139	0.0028
		RRMSE	0.1284	0.0231	0.0149
	(100, 200)	RB	-0.0573	-0.0018	0.0091
		RRMSE	0.0593	0.0124	0.0147

$\rho_{exact} = 0.8971$

(3, 6.2)	(10, 10)	RB	-0.4517	-0.0841	-0.0069
		RRMSE	0.4795	0.1588	0.1152
	(20, 30)	RB	-0.2603	-0.0077	0.0177
		RRMSE	0.2739	0.0553	0.0628
	(50, 50)	RB	-0.1798	-0.0174	0.0152
		RRMSE	0.2002	0.0570	0.0542
	(100, 200)	RB	-0.0682	0.0035	0.0185
		RRMSE	0.0738	0.0287	0.0312

$\rho_{exact} = 0.7581$

(3, 10.3)	(10, 10)	RB	-0.5063	-0.0508	0.0516
		RRMSE	0.5541	0.1758	0.1508
	(20, 30)	RB	-0.3423	-0.0507	0.0157
		RRMSE	0.3634	0.1070	0.1244
	(50, 50)	RB	-0.2718	-0.0313	0.0199
		RRMSE	0.2810	0.0787	0.0688
	(100, 200)	RB	-0.1016	0.0067	0.0417
		RRMSE	0.1121	0.0434	0.0621

$\rho_{exact} = 0.5679$

(3, 20.4)	(10, 10)	RB	-0.6922	-0.0509	-0.0238
		RRMSE	0.7190	0.2917	0.2629
	(20, 30)	RB	-0.4200	-0.0654	-0.0138
		RRMSE	0.4661	0.2100	0.2258
	(50, 50)	RB	-0.3587	-0.0549	0.0391
		RRMSE	0.3802	0.1210	0.1404
	(100, 200)	RB	-0.1744	0.0120	0.0494
		RRMSE	0.1879	0.0715	0.0917

Table 3. The RB and RRMSE of the three estimators $\hat{\rho}_k$, $\hat{\rho}_{p_1}$ and $\hat{\rho}_{p_2}$ when the data are simulated from pair Weibull distributions with equal shape parameters ($\beta_1 = \beta_2 = 3$).

$\rho_{exact} = 0.9637$

(α_1, α_2)	(n_1, n_2)		$\hat{\rho}_k$	$\hat{\rho}_{Al}$	$\hat{\rho}_3$	$\hat{\rho}_{P2}$
(1, 1.2)	(10, 10)	RB	-0.3857	-0.0874	-0.1306	-0.1041
		RRMSE	0.4352	0.1414	0.1932	0.1692
	(20, 30)	RB	-0.1682	-0.0124	-0.0297	-0.0139
		RRMSE	0.1868	0.0393	0.0562	0.0404
	(50, 50)	RB	-0.1075	-0.0179	-0.0234	-0.0199
		RRMSE	0.1171	0.0392	0.0438	0.0400
	(100, 200)	RB	-0.0521	-0.0017	-0.0039	-0.0015
		RRMSE	0.0545	0.0135	0.0130	0.0135

$\rho_{exact} = 0.8398$

(1, 1.5)	(10, 10)	RB	-0.3222	-0.0482	-0.0798	-0.0549
		RRMSE	0.3857	0.1739	0.1948	0.1878
	(20, 30)	RB	-0.2104	-0.0563	-0.0642	-0.0605
		RRMSE	0.2260	0.0905	0.0983	0.0948
	(50, 50)	RB	-0.1113	-0.0173	-0.0225	-0.0106
		RRMSE	0.2023	0.1106	0.1182	0.1249
	(100, 200)	RB	0.0744	-0.0104	-0.0168	-0.0163
		RRMSE	0.0879	0.0419	0.0473	0.0478

$\rho_{exact} = 0.7069$

(1, 1.8)	(10, 10)	RB	-0.3070	-0.0146	-0.0671	-0.0344
		RRMSE	0.3490	0.1525	0.1989	0.1952
	(20, 30)	RB	-0.1657	0.0115	-0.0091	0.0105
		RRMSE	0.2241	0.1182	0.1177	0.1190
	(50, 50)	RB	-0.1306	-0.0059	-0.0209	-0.0166
		RRMSE	0.1754	0.1000	0.0990	0.1003
	(100, 200)	RB	-0.0638	-0.0038	-0.0017	-0.0003
		RRMSE	0.0800	0.0456	0.0424	0.0424

$\rho_{exact} = 0.2984$

(1, 3.5)	(10, 10)	RB	-0.2619	-0.0988	-0.1227	-0.0538
		RRMSE	0.4741	0.2931	0.4394	0.4566
	(20, 30)	RB	-0.1373	-0.0267	-0.0178	0.0082
		RRMSE	0.3179	0.2201	0.3010	0.3128
	(50, 50)	RB	-0.1275	-0.0409	-0.0267	-0.0018
		RRMSE	0.2786	0.1854	0.2659	0.2781
	(100, 200)	RB	-0.0694	0.0023	0.0131	0.0165
		RRMSE	0.1361	0.0955	0.1131	0.1126

Table 4. The RB and RRMSE of the three estimators $\hat{\rho}_k$ and $\hat{\rho}_P$ when the data are simulated from pair Weibull distributions with different scale and different shape.

$\rho_{exact} = 0.9793$

(α_1, α_2)	(β_1, β_2)	(n_1, n_2)		$\hat{\rho}_k$	$\hat{\rho}_3$
(1, 1.2)	(2, 1.8)	(10, 10)	RB	-0.2213	-0.0704
			RRMSE	0.2529	0.1043
		(20, 30)	RB	-0.1433	-0.0387
			RRMSE	0.1700	0.0751
		(50, 50)	RB	-0.0535	-0.0030
			RRMSE	0.0651	0.0203
		(100, 200)	RB	-0.0325	-0.0048
			RRMSE	0.0362	0.0162

$\rho_{exact} = 0.8435$

(1, 1.5)	(3, 1.9)	(10, 10)	RB	-0.2120	-0.0117
			RRMSE	0.2700	0.1280
		(20, 30)	RB	-0.1419	-0.0197
			RRMSE	0.1796	0.0906
		(50, 50)	RB	-0.0993	-0.0253
			RRMSE	0.1158	0.0556
		(100, 200)	RB	-0.0614	-0.0158
			RRMSE	0.0688	0.0356

$\rho_{exact} = 0.6893$

(1, 1.8)	(4, 2.1)	(10, 10)	RB	-0.3760	-0.1322
			RRMSE	0.4292	0.2403
		(20, 30)	RB	-0.1624	-0.0148
			RRMSE	0.2064	0.1413
		(50, 50)	RB	-0.1166	-0.0227
			RRMSE	0.1542	0.1126
		(100, 200)	RB	-0.0783	-0.0033
			RRMSE	0.0987	0.0561

$\rho_{exact} = 0.3929$

(1, 3)	(6, 2)	(10, 10)	RB	-0.4320	-0.1404
			RRMSE	0.5979	0.4086
		(20, 30)	RB	-0.1983	0.0259
			RRMSE	0.3472	0.2311
		(50, 50)	RB	-0.1616	0.0006
			RRMSE	0.2437	0.1836
		(100, 200)	RB	-0.1087	-0.0160
			RRMSE	0.1350	0.0758

References

Alodat, M. T., Al Fayez, M. & Eidous, O. (2022). On the asymptotic distribution of Matusita's overlapping measure. *Communications in Statistics - Theory and Methods*, 51(20), 6963-6977.

Al-Saidy O, Samawi H. M. & Al-Saleh M. F. (2005). Inference on overlapping coefficients under the Weibull distribution: Equal shape parameter. *ESAIM: Probability and Statistics*, 9, 206-219.

Dhaker, H., Ngom, P. & Mbodj, M. (2019). Overlap coefficients based on Kullback-Leibler divergence: exponential populations case. *International Journal of Applied Mathematical Research,* 6(4), 135-140.

Eidous, O. & Al-Daradkeh, S. (2022). Estimation of Matusita overlapping coefficient $\boldsymbol{\rho}$ for pair normal distributions. *Jordan Journal of Mathematics and Statistics (JJMS)*, 15(4B), 1137-1151.

Eidous, O. & Al-Talafha, S. (2022). Kernel method for overlapping coefficients estimation. *Communications in Statistics - Simulation and Computation*, 51(9), 5139-51516.

Eidous, O. & Shourrman, A. (2022). Numerical integral approximation to estimate Matusita overlapping coefficient for normal distributions. *Journal of Mathematical Techniques and Computational Mathematics*, 1(3), 264-270.

Federer, W., Powers, L. & Payne, M. (1963). Studies on statistical procedures applied to chemical genetic data from sugar beets. *Technical Bulletin, Agricultural Experimentation Station*, Colorado State University.

Gastwirth, J. (1975). Statistical measures of earnings differentials. *The American Statistician*, 29(1), 32-35.

Ichikawa, M. (1993). A meaning of the overlap pedarea under probability density curves of stress and strength. *Reliability Engineering and System Safety*, 41(2), 203-204.

Inman, H. & Bradley, E. (1989). The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. *Communications in Statistics - Theory Methods,* 18, 3851-3874.

Matusita, K. (1955). Decision rules based on the distance for problem of fir, two samples, and estimation. *Ann. Math. Statist*, 26, 631-640.

Mulekar, M. & Mishra, S. (1994). Overlap coefficients of two normal densities: equal means case. *Journal of the Japan Statistical Society*, 24(2), 169-180.

Pastore, M. (2018). Overlapping: a R package for estimating overlapping in empirical distributions. *Journal of Open Source Software*, 3(32), 1023.
Pianka, E. (1973). The structure of lizard communities. *Annual review of ecology and systematics*, 4(1), 53-74.

Samawi, H. M. & Al-Saleh, M. F. (2008). Inference of overlapping coefficients in two exponential populations using ranked set sample. *Communication of Korean of Statistical Society*, 15(2), 147-159.